THE RUDIMENTS OF RELATIVITY

Lectures delivered under the auspices of the University College, Johannesburg, Scientific Society.

BY

JOHN P. DALTON,
M.A., D.Sc.,
Fellow of the Royal Society of South Africa,
Professor of Mathematics at University College, Johannesburg.

PUBLISHED BY THE
COUNCIL OF EDUCATION, WITWATERSRAND, JOHANNESBURG.
1921.

Price Five Shillings.

THE RUDIMENTS OF RELATIVITY

Lectures delivered under the auspices of the University College, Johannesburg, Scientific Society.

BY

JOHN P. DALTON,

M.A., D.Sc.,

Fellow of the Royal Society of South Africa,
Professor of Mathematics at University College, Johannesburg.

PUBLISHED BY THE

Council of Education, Witwatersrand, Johannesburg.

1921.

Price Five Shillings.

PREFACE

The following pages contain four lectures which were prepared as a Presidential address to the Scientific Society of this College. Although their preparation entailed a considerable amount of labour I regarded it as a normal incident of academic life, and had no intention of publishing my views. Many of my colleagues, hearing of the topic proposed for discussion, expressed a desire to be present, and the Council of the Scientific Society accordingly decided to throw the lectures open to the public. The response to that invitation was astonishing and stimulating; it showed that even in what is regarded (quite unjustly, I think) as a very frivolous and material centre there is a widespread desire for intellectual stimulus. Various members of that audience have asked me to publish my lectures; I therefore print them now substantially as delivered, although I should have preferred a less crude and incomplete form of presentation.

It will be borne in mind, I trust, that I was primarily addressing students whose scientific knowledge and reading could hardly be called extensive; it was therefore essential that I should divest the theory of its customary mathematical elegance, and present instead its physical basis and implications. The result is an absence of

precision that is hardly commendable, but perhaps the attempt at definite physical interpretation will afford a suitable introduction to a mathematical treatment.

I must express my cordial thanks to the Council of Education, Witwatersrand, for undertaking the publication of these lectures. It is largely owing to their consistent zeal and financial support that education in the Transvaal is at present so highly developed, and the fact that their activities, as in the present instance, are by no means confined to utilitarian purposes is of happy augury for the future intellectual life of the community.

JOHN P. DALTON.

Johannesburg, August 1st, 1921.

CONTENTS.

"*In creative thought common sense is a bad master.*"

A. N. WHITEHEAD.—Introduction to Mathematics.

FIRST LECTURE.

THE BACKGROUND.

SECTION 1.—TIME-ORDER.

The object of the investigation of natural phenomena, which is the purpose of Natural Philosophy, is the introduction of order into the apparently chaotic appearances of Nature. In accomplishing this, reliance is placed upon two fundamental *a priori* relations of order. The first of these relational processes is an immediate application of the serial order involved in the idea expressed by the words "before" and "after." Of two events, A and B, which present themselves to my consciousness, I can generally state that A either precedes or succeeds B. To avoid complications, I consider only events of sufficiently short duration; otherwise there might be overlapping, and it would then be possible for event A to both precede and succeed event B. If, for example, event A were my addressing you to-night, and event B were the passage of the hands of the clock across the position usually read as 9 o'clock, then A would both precede and succeed B. We shall exclude such overlapping, and imagine that we are dealing with events of such short duration that we can relate them in definite serial order. This intuitive seriation yields us a time order in which the conscious mind can arrange its experiences.

The process is not so simple as I seem to indicate. Without the intervention of Memory the time ordering relation could not operate; the present would vanish from consciousness instantaneously; no temporal seriation of events would be possible. I just take the relation as an

acknowledged primal intuitional process, however complex it may be psychologically, and make no further attempt to analyse it.

The fact that time-order is intuitional leads to difficulties. Since it is intuitional, it is necessarily individual. There is then no guarantee that a sequence of events *a*, *b*, *c*, will present themselves in the same order to two distinct individuals. Consider, say, the discharge of a gun and the subsequent burst of the shell. A man near the target will hear the burst of the shell before he hears the report of the discharge. A man near the gun will hear the discharge before he hears the burst. And, moreover, there will be a line of positions between the gun and the target, such that to observers stationed there, discharge and burst would appear simultaneous. How then are we to account for these differences in the observed sequence of events? It would hardly be satisfactory to postulate vital differences in the mental structure of the individuals concerned, for then no coherent description of our common experiences would be possible. We must assume that the time-relational processes of the observers are identical, and that the apparent discrepancies of time-sequence originate in the circumstances of propagation. In this example we harmonize the different descriptions by ascribing finite velocities of travel through the air to the shell and to the sound, that of the shell being the higher.

Let us now consider an observer who is so placed that he hears discharge and burst simultaneously. The sounds are carried by the air. If there is no wind his position will be somewhere on a line at right angles to the join of gun and target, somewhat nearer the target, because the report of the discharge gets a start on that of the burst on account of the finite velocity of the shell. If a wind now blows across the field from the gun to the

target the discharge and the burst will no longer seem simultaneous. Although the relative positions of the gun, target and observer remain unaltered, the observed time seriation of the occurrences is different. The observer adheres to his postulate of the constancy of his time-ordering relation, and accounts for the observed variation by the fact that the air particles are themselves in motion towards him from the gun. Hence the sound of the discharge travels towards him more quickly, and that of the burst more slowly, than would be the case in still air. It is important to realize that this harmonization is rendered possible by the known existence of the material medium of propagation—the air; we could not very well offer that explanation if the mechanism of sound propagation were unknown.

SECTION 2.—THE TIME-INTERVAL.

This somewhat vague time order does not suffice. We are not content with knowing that of three events A, B and C, A preceded B and B preceded C; we are also interested in the relative values of the intervals between A and B and between B and C; we need not only a time sequence, but also a time-interval scale.

Now a time scale cannot be intuitional; it must be conventional. Any attempt to depend upon intuitional quantitative time estimates would be invalidated by psychological complications; for the apparent remoteness or proximity of an event depends not only upon the date of its occurrence, but also upon the intensity of the feelings it provoked. Quantitative time values arise only from comparison with some arbitrary series of events whose apparent permanence and regularity give it the stability essential to a reliable standard of comparison.

Natural periodicities provide us with an ordered sequence against which we may compare our experiences

in so far as their durations are concerned. The particular periodicity selected by civilized peoples is the apparent diurnal rotation of the heavens. As we regulate our lives by the sun, it is natural to take its regular passage across the sky as affording a suitable time scale. Consider an individual at a certain place who is able to observe the repeated passages of the sun across the meridian; the interval between successive transits provides him with a unit time interval. He may move about the earth in any way, quite satisfied with his time scale, and finding no discrepancies in his experience as long as he is intent upon his own doings only. Should he, however, desire to compare notes with another individual whose time scale is based upon the same periodicity he will experience difficulty of correlation. Should he, for instance, have gone round the world, while the other had remained at the initial point, they would find a difference of a unit in their accounts of the interval that had elapsed since their last meeting. Our observer might attempt a correlation by postulating erratic celestial movements depending upon the part of the earth from which they were observed. That, however, would hardly commend itself to him, for it would make the description of occurrences on the earth a very complicated process. He would reason rather that as his time unit depended upon the motion of the sun relative to himself, his own movements must be taken into account if he is to attain consistency.

From this it is evident that the time of the physicist and of the philosopher, an ordered succession of instants, or, in the words of Newton, "an absolute true and mathematical time, which in itself, and of its own nature, flows equally," is not by any means a direct datum; it is not disclosed to us in our sense-perceptions. It is an intellectual construction attained by processes of correla-

tion and of abstraction. The process of correlation may involve a consideration of the physical circumstances of a material medium, and it certainly does involve a consideration of relative motion.

SECTION 3.—DIRECTION ORDER.

The second fundamental intuitional process which enables us to order our experiences in the world of sense we may call the direction-ordering relation. We automatically recognize the difference between "here" and "there." I am here, you are there; Johannesburg is here, Cape Town is there; the Earth is here, the Sun is there. We see at once that it is only relative position that is apprehended in our sense perception. If you are asked over the phone "Where are you?" and if you answer "I am here," the reply you get will (should your questioner's vocabulary be at all adequate), enlighten you upon the essential relativity of position. If I consider only my own immediate field of sense perception, I can discriminate between different directions relative to myself. The window is on my right; the door is on my left; the blackboard is behind me; the ceiling is above me. If I am to get a consistent scheme of things I must assume that the same directional discriminative sense will be exercised by others. Its direct application, however, leads to divergent descriptions. The window is on your left, the door is on your right; the ceiling is above you, the blackboard is in front of you. We agree in some particulars, we differ in others. The difference originates in the relativity of our descriptions due to the individualness of our perceptions. We endeavour to correlate our different experiences, to throw our various data into a common stock, and extract therefrom a method of description with which we can all agree. A suitable harmonization could be effected by agreeing to the fixation of

some position as a standard position of reference, deliberately imagining ourselves to occupy that position. Let us therefore imagine ourselves over in that corner, with our back to the door. Then the window is on our right, the blackboard on our left. But *our* right and *our* left are no longer the directions disclosed to us by our sense perceptions functioning in our actual positions; they are intellectual constructions arrived at by the deliberate suppression of immediate sense data, and an imagined removal of our system of reference from the actual to a supposed position. We sacrifice immediacy of perception in order to gain sameness of description. This sameness of description is a very important feature of relativity discussions; in technical language we speak of a description which is the same for all members of a group of observers as being *invariant* for that group.

Section 4.—The Distance-interval.

We have not yet, however, attained a method of giving an account of position. We can agree upon directions in a conventional system of reference, but there is more than that involved. Going back to our corner we find that the first window is on our right, so also is the second; and yet they are not in the same place. We recognize that if we were to proceed from our corner in the specified direction we should encounter those windows at different times. In any given direction we can arrange the things we encounter in serial order, just as we arrange our experiences in a temporal series. But this new seriation is not simply intuitive, for, having fixed a direction how do we determine distances along it? We must either imagine ourselves proceeding along the direction indicated, and encountering the objects located therein in a definite order; that is time seriation complicated by the concept of motion. Or we may measure off the distances.

But measurement entails successive applications of an arbitrary standard of length, and an enumeration of the number of applications before reaching the specified object; the time-ordering relation is still there.

Section 5.—The Interconnection of Distance and Time.

To my own consciousness the time ordering relation and the direction ordering relation are distinct and independent processes; but they are individual. In any theory of physical knowledge the accounts of all possible observers must be synthesized. We must assume that other observers are human beings like ourselves, endowed with the same fundamental order intuitions. Nevertheless we find from experience that their descriptions of observed events do not always agree with ours, and the correlation of such divergent descriptions we call a *transformation* from one point of view to another. If every description changed with every transformation the world would be entirely individual and subjective; objective reality is rendered possible by the discovery of permanences which are unaffected by alterations in the point of view. Such permanences are the invariants of the transformations.

Order relations do not suffice for the correlation of my own experience *here* and *now* with my own, or some other person's, experience *there* and *then*; in such cases the necessity for *intervals* of time and *intervals* of distance becomes urgent. But our interval of time is defined explicitly in terms of motion, which involves distance; and our interval of distance is defined either explicitly or implicitly in terms of motion, which involves time; distance-interval and time-interval are thus interwoven in the concept of motion. Of these three we may arbitrarily fix upon any two, and the third is then given

by their interrelation. Hitherto we had always fixed upon unit time and unit distance as fundamental, and velocity was then determined as a derived unit. But we are only given the relation between the three, and may start in any way we choose. Einstein proposes as fundamental a unit velocity (which is the ratio of distance to time), and lets time and distance units then adjust themselves. As the units are conventional and arbitrary, there is no logical reason why such procedure should not be adopted; indeed, in taking velocity as a starting point we explicitly recognize the inextricable interconnection of our units of time and of distance; but we must not feel surprised if different invariants prove characteristic of different methods of description.

SECTION 6.—THE PHYSICIST'S DESCRIPTIVE ELEMENTS.

Before discussing Einstein's theory, however, we must examine the traditional views a little more closely.

In formulating his scheme for describing relative position the physicist suppresses the fundamental intuition of direction, and emphasizes the derivative concept of serial linear order. He fixes his centre of reference in any convenient position and takes three directions, to the right, say, in front and vertically upwards. He then specifies the position of a place, with reference to these axes, by numbering its co-ordinates; that is to say, by specifying the distances one must go to the right, in front, and up, in order to arrive at the given point. No one could consider this a direct positional intuition. But the object of the physicist is to measure and this system of reference gives him an extremely convenient frame into which he can fit his observations. The convenience arises from the fact that he may put his centre of reference anywhere he likes, in the laboratory, in the street, at the centre of the earth, in the sun; he may choose his initial

direction of reference in any way; the co-ordinates of any two points will naturally depend upon his arbitrary choice, but their relative distance will always be calculated in exactly the same way from the co-ordinates no matter what their numerical values may be. Now relative distances determine the shapes of bodies; hence the shapes of bodies will be invariant; they will be independent of the particular frame of reference selected. That, you will agree, helps considerably in affording coherent and consistent descriptions. It does not give us the absolute space, which, in the words of Newton, "by virtue of its own nature, and without reference to any external object, always remains the same and is immovable." On the contrary, it is purely empirical, but having the distinct advantage of retaining the rigidity of bodies which to our senses remain substantially unaltered no matter how we place them.

In experience we encounter nothing like the physicist's point in space. I speak of here, at this point, and put my finger on the spot. But where my finger touches the table is not a point; it is a patch of more or less indefinite outline, depending on which finger I put down; and on how hard I press with it. Were I to point with a needle instead of with my finger I should still get a patch of some extension which could be seen and measured under a microscope.

Even so is it with time. We know of no instants as such. You may think that the only temporal reality is the present instant *NOW*; for the past is gone; the future is still to come. But we are conscious of time only through the occurrence of events, and we can know of no events that are instantaneous. We speak of instantaneous photography, but the quickest exposure there known is of definite duration, some measurable fraction of a second. Not only is the instant of the physicist unrevealed to us

in sense-perception, but even our actual present "now" is blurred and indistinct. Memory retains the immediate past; the proximate future can be in part anticipated. As I speak, even though my words go in at one ear and out at the other, at least two or three of them will be present to your consciousness simultaneously although uttered in succession; while, with a little attention, you can often anticipate the conclusion of a sentence before I have completed its enunciation. Our places are never points; our times never instants; there is always more or less indefinite extension or duration. This discrepancy between our experiences and our descriptions of them is not, however, a very serious difficulty. The abstractive processes whereby "points" and "instants" are constructed are all-important in mathematics, but I do not propose to discuss them here; we shall take them for granted.

But the constructed world of physics and the actual world of sense differ in other respects. The physicist keeps rigorously apart his two sets of descriptive elements. He busies himself with what is happening in the neighbourhood of a specified point at a definite instant, or with what is happening during a specified duration of time at a definite point; he effects a neat and precise cleavage between the two concepts. In the world of events no such cleavage exists. I think, say, of myself now, and myself a year ago; surely that is a pure time interval. On the contrary. Now, I am here in this room in Johannesburg; a year ago I was in Scotland. The interval between the two events, myself now, and myself then, involves a time lapse of a year, and a distance change of some 5,000 miles, in each case relative to a system of reference attached to the earth. The intervals between two events in Nature are always of this type. The historian thinks of the growth of British South Africa within a century; the

thought embraces growth in extent and growth in time. There is no difficulty apparently in giving adequate and comprehensive verbal expression to the idea; it is not considered in any way strange. But when the mathematical relativist endeavours to give precise and formal expression to the same concept, when he endeavours to measure the interval between two events taking all the observed quantities into account, and tries to find a formula which will yield the same description of that interval for all possible observers of the events, then he is relegated by sane and sensible people to the dismal wilderness of the fourth dimension.

The outcome of this preliminary discussion is that what we are most conscious of in Nature is change, and change entails motion. Motion at a durationless instant is meaningless; we cannot imagine a state of change without taking into account the immediate past and the immediate future. Therefore in giving an account of the changes we observe we have to resort to relational processes. Two of the simplest of these we have just considered; they are examples of what are called in logic two-termed relations. You cannot have a one-termed relation. The words "there" and "then" are of themselves of no significance; we must be supplied with the other term of the relation before any meaning emerges. But the essentially distinct natures of these two relations are seriously modified by the methods of defining intervals of relationship. Owing to the fact that our basic data are motional events we cannot obtain independent standards.

We shall, however, accept for the present the familiar point space and instantaneous time of the physicist as intellectual constructions affording us a convenient system of description, but we do so with the reserve that they are derivative concepts, and in no way primitive sense data.

Section 7.—Newton's Laws.

We may now turn to a consideration of the procedure adopted in the investigation of natural phenomena. It originates in the belief that natural processes are inherently and essentially harmonious. It generally entails

First, observation of a series of facts;

Secondly, the formulation of a theory or hypothesis which would adequately cover these facts;

Thirdly, the deduction, by purely mental processes, of further consequences arising out of such hypotheses; and

Fourthly, the verification of such further consequences by observation and experiment.

As long as we find such further conclusions concord with the results of observation, so long are we satisfied with our theory, and we speak of it as a law. There is nothing irrefragable about such a law, for new facts may at any time come to light necessitating its modification or even its abandonment. Such laws harmonize our discrete experiences of Nature co-ordinating them into homogeneous groups on the basis of assumptions of some inherent probability; they enable us to reason coherently about our physical observational data. It is in this sense, and in this sense only, that we speak of such hypotheses as Laws of Nature.

Amongst such laws of Nature is a group which lies at the root of all relativity discussions—the laws of motion enunciated by Newton about the middle of the seventeenth century. We have seen that motion is a primal fact in nature. It had been studied from the earliest times. Philosophers had discussed its objectivity; experimentalists had measured it. Newton collected the existing knowledge under the heading of three great Principles. These are:

I. *The Principle of Inertia*, by virtue of which a body at rest remains at rest, or if in motion, keeps moving uniformly in a straight line, unless constrained by some external agency to change that state of rest or of uniform motion. We say that matter is inert; it cannot of itself initiate or alter motion.

II. *The Principle of Acceleration*. When a state of rest or of uniform rectilinear motion is being changed, the force operative is proportional to the inertia of the moving body and also to the rate at which its velocity is changing. This rate of change of velocity is called the acceleration of the body.

III. *The Principle of Reaction*. To every action there is an equal and opposite reaction. This principle asserts that every action between bodies is of the nature of a stress. The most important illustration of it that we shall encounter arises in problems of rotation. If you whirl round a stone on a string you must provide a constant pull in order to keep the direction of motion constantly changing; that is in accordance with the principle of inertia. But the stone, by virtue of its motion, exerts an equal and opposite pull on you. We call such a force, existing only through motion, an inertial force, in particular, if the motion is one of rotation, a centrifugal force; and we say that the string is in a state of stress under the influence of the pull you exert on the stone and the kinetic reaction of the stone. Were the stone not in motion the same state of affairs could be maintained, in so far as the forces are concerned, by pulling the stone outwards with the same force as that with which you pull it inwards. But it is of importance to notice that such a system could not produce or alter motion.

For our purpose these laws of motion must be supplemented by a fourth law, also due to Newton, namely:

IV. *The Principle of Gravitation.*

About 50 years before Newton's time, extraordinary perseverance with arithmetical computation had led Kepler to the discovery that all planets move round the sun in elliptical orbits. The phenomenon of gravitation as exemplified by falling bodies had also been known for some considerable time. Astronomers were then concerned with the relationship, if any, existing between the planetary movements and gravitation. The synthesis was effected by Newton. According to his theory the activity that causes the fall of bodies near the surface of the earth is characteristic of matter generally; its intensity depends upon the mass of the body in which it originates; it spreads throughout the neighbourhood of that body according to the natural law of spread for Euclidean space, namely, that the intensity of the activity diminishes according to the square of the distance from its origin.

Some points in connection with Newton's laws had better be noticed at this stage. The law of inertia mentions a body at rest. What do we mean by "at rest"? If I stand still a moment am I not at rest? I am, relatively to the earth. But the earth is carrying me round its own axis and round the sun at what I should call considerable speed. Rest is therefore a relative term; it becomes significant only when referred to some system of reference. In Newton's scheme that system is his absolute space.

Again, a body in motion if left to itself keeps on moving in a straight line. What then is a straight line? Our direction ordering intuition gives us a straight line in our own immediate neighbourhood; we *see* that a number of points are in the same direction from us, and we say that they fix a straight line. Our intuitional straight line is then the path of a ray of light. But suppose we look at

something at the bottom of a basin of water; we see it in a certain direction. Let us move what seems to us to be a straight rod in that direction; it will miss the object. Our visual direction is no longer a straight line. It may well be then that a light ray coming to us from afar is curved without our knowing it.

We may try defining our straight line as the shortest distance between two points. How are we to find the shortest distance? Measure it. Will it matter if our measuring rod is bent? Surely so. How then can we tell if our rod is not bent? Apparently we cannot tell until we know what a straight line is.

We shall have to examine these matters more closely later on. In the meantime all we can do is to accept the principle of inertia as defining a straight line; that is to say, a straight line is the path of a moving body which is sufficiently far removed from all other bodies. The definition then becomes purely conceptual, because we can never get quite away from other influences; at any place that we can reach, our moving body will be influenced by gravitation; we can never leave the gravitational field. The first law of motion, therefore, instead of being an expression of experienced fact, becomes a matter of hypothetical definition.

In spite of this and other difficulties involved in Newton's mechanical system, it has more than justified itself by the verification of deductions based upon it. It has welded such diverse phenomena as the planetary motions, the fall of bodies, the swing of a pendulum, the rise and fall of tides and the precession of the earth's axis into a compact and harmonious dynamical system. Relativists claim that certain physical facts are now known which do not fit naturally into this great scheme, and they therefore propose to modify it. The modifications are very minute numerically, but are of vast philosophical import.

SECTION 8.—NEWTONIAN RELATIVITY.

We are now in a position to appreciate, and accept, the doctrine of mechanical relativity, which issues from the laws of motion. The doctrine was explicitly stated by Newton in the fifth corollary to his laws of motion:—"The motions of bodies within a given space are the same amongst themselves, whether the space is at rest, or moves uniformly forwards in a straight line." He instances the case of a ship where, he says, "all motions happen in the same manner whether the ship is at rest, or is carried forward in a right line." As long as the motion of the ship is steady and uniform our life proceeds as usual; the principles of inertia, of acceleration, of reaction obtain just as on shore. No observation or experiment carried out within the ship itself could reveal its motion through the sea. Of course if you may experiment upon the sea as well, that is, the material medium through which the ship is being conveyed, then the relative motion of the ship can be detected. If, for instance, you throw a cork into the sea, you will see it fall behind.

As soon, however, as a rough sea disturbs the uniform motion of the ship, we take a poignant interest in the failure of the relativity principle; we notice that things do not go on as if the ship were at rest; we are particularly conscious of the fact that action and reaction are no longer perfectly compensated, and the normal state of stress becomes one of distress. You can experiment on a small scale with a weight attached to a spring balance. If you move the balance steadily up or down, the reading will be constant during the motion; the weight of the object and the tension of the spring are in compensation. If, however, you move it about in jerky fashion the lack of compensation will be evidenced by the correspondingly erratic readings of the pointer.

Newtonian relativity is therefore applicable only to the case of systems which move uniformly and rectilinearly. Such systems are called, for short, inertial systems. Newtonian relativity may then be stated thus :—That the laws of motion are identical for all inertial systems.

To appreciate the significance of this statement let us imagine two trains moving at different constant speeds along straight parallel tracks. Let each train carry an experimenter, A and B, with a fully equipped mechanical laboratory, and provide them with wireless telephones so that they may communicate with each other without disturbing their motion in any way. Each, within his own compartment, will find that bodies thrown in the air describe parabolas, that simple pendulums to which they ascribe the same lengths have the same time of swing, that for the same bodies similar spring balances give the same readings, that identical tuning forks emit the same note—they will find, in short, that the laws of motion and all their consequences are accurately obeyed. That is not surprising, for the laws of motion deal with accelerations, and there is no relative acceleration between the two trains, and therefore no reason why the laws should be other than identical. The principle of Newtonian or mechanical relativity then amounts to this, that as long as our observers concern themselves only with their respective spheres of experiment, their results will be in perfect accord.

Section 9.—The Newtonian Transformations.

As soon, however, as they look about them and observe things outside their own particular systems, then numerical discrepancies may well arise. To fix our ideas suppose A's train is moving North at 20 m.p.h. and B's South at 15 m.p.h. If each looks at the other in passing, each will state that the other is travelling in a straight

line at constant speed; but A will say that he is at rest and that B is going south at 35 m.p.h. whereas B will say that A is going north at 35 m.p.h.

Again, if a bird is flying towards the railway track from the East, say, as seen by a man standing beside the track, A, going north, will say that it is flying towards him from the north-east with a certain speed, while B, going south, will insist that it is flying towards him from the south-east with a different speed. No doubt A and B will feel embarrassed by the discordant results of their observations, and will endeavour to attain a common standpoint of description. A may accordingly say to B, "You give everything you see an added speed of 35 m.p.h. in the direction in which you are going and you will see things as I see them." And if B is at all human, he will answer: "Why should *I* bother; it is *you* who are moving. You give everything you see a velocity of 35 m.p.h. in the direction you are moving, and you will see things as I see them." As long as they restrict their observations to the two trains neither A nor B will admit that he is moving, so neither will see any reason for altering his observed data in the way suggested by the other. A, however, will recognize that if he imagines the modifications he suggested carried out on B's reported observations, he will make them identical with his own; and B will see that he can do the same thing. This modification constitutes a transformation from one point of view to another. The particular transformation which we are at present discussing enables us to state in what respect things will differ when observed from various inertial systems. It is called the Newtonian or Galilean transformation. In carrying out this transformation we are going a step further than the principle of mechanical relativity as stated by Newton. All that Newton's principle amounts to is this: that within any uniformly moving system,

things go on in the same way no matter how fast or how slowly it may be moving. Now, however, we are considering more than one such system; we are busying ourselves not only with what is going on in each system by itself, but we are also looking into other systems which are moving relatively to our own, and trying to describe what is happening there. We find in such case that the laws of motion hold for all such systems, but that some relations derived from these laws, such as velocities, differ numerically according to the system from which they are observed.

This non-invariance of velocity with respect to different inertial systems is of some importance, because the hypothesis of the invariance of the velocity of light for all observers is a basic assumption of Einstein's physical theory. In his elementary exposition[1] Einstein supposes a railway carriage moving along a straight track with constant speed, v, relative to the embankment, and imagines a ray of light sent along the embankment in the same direction with speed c. If we were to treat the ray of light as a movement of air, say, then the velocity the light relative to the carriage would be $c - v$. "The result," he says,[2] "comes into conflict with the Principle of Relativity set forth in Section V. (that is, the Newtonian, or mechanical principle that we are just now discussing). For like every other general law of Nature, the law of transmission of light in vacuo must, according to the principle of relativity, be the same for the railway carriage as reference body as when the rails are the body of reference." He proceeds to emphasize this[3] :—"If every ray of light is propagated relative to the embankment with the velocity c, then .... it would appear that

---

[1] *Einstein*: Relativity. London, (1920).

[2] *op. cit.* p. 18.

[3] *op. cit.* p. 19.

another law of propagation of light must necessarily hold with respect to the carriage—a result contradictory to the principle of relativity."

We have seen that on the Newtonian system velocities must, and do, vary according to the inertial system from which they are observed. There seems therefore no contradiction as stated by Einstein. We shall have to return to this point in the next lecture, and find some other reason for treating the velocity of light as something unique; it is the crucial question in physical relativity.

Another question of comparison arises. A and B agree about their forces and accelerations; will they agree about the shapes of things they see? Suppose B has a spherical ball in his compartment and that A in passing looks at it; will it be a spherical ball to him? and will it have the same diameter? To compare their respective points of view we shall have to apply the appropriate transformation. We have seen that if A wants to see things as B sees them he must give everything a constant velocity of 35 m.p.h. That will apply to every point of the ball, so that the relative distances between the different parts of it will remain the same. There will therefore be no difference between their descriptions of the shapes and sizes of things they see. Rigid bodies will accordingly preserve their dimensions unaltered whether they are moving with the observer or not, and consequently a geometry based upon congruence will hold for all inertial systems.

It follows that we can find no one inertial system of reference such that descriptions of observations within and without the system are simpler with respect to it than with reference to any other such system. Did we find any such specially favoured system we should naturally regard it as fundamental, and should attribute to it an absolute significance. But the Newtonian trans-

formations offer no such unique preference; all inertial systems are equivalent for the description of all observed facts.

Section 10.—Accelerated Systems and the Principle of Equivalence.

As long as we are concerned with purely terrestrial phenomena it is convenient to fix our axes of reference in the earth. The earth seems to our ordinary senses to be stable and motionless, and we detect the motion of our trains or other bodies by observing a point fixed on the surface of the earth. But when we extend this limited field of experience, we see that the earth is not an inertial system at all, and axes attached to it are in rotation with respect to the sun. We must be able to appreciate the difference between an inertial system and a rotating system, as this forms the basis of the later generalisations of relativity. The difference is essentially this: If we perform experiments within a rotating system, we can observe results which *may* be ascribed to the motion of rotation, whereas, as we have seen, within inertial systems no experiment whatever can reveal the motion of translation. You will notice I say *may* be ascribed to the rotation; the basis of the general relativity theory is that they need not be so ascribed, but could originate in other ways. In the meantime we must record the facts. Here I have some flexible strip, arranged in approximately spherical shape; when I set it rotating, you see that it is no longer spherical. The Earth in the same way bulges about its equator, the bulge being generally attributed to its rotation about its axis. Usually the relative rotation of the earth is brought home to us by observing the revolutions of celestial bodies, and if we were concerned with such movements only, it would be a matter of indifference whether we regarded the heavens

as still and the earth rotating on its axis, or the earth as still and the celestial bodies revolving in tremendous orbits with correspondingly huge velocities. But here we are concerned with more than relative motion. Suppose the earth's atmosphere were quite opaque, we could still measure its surface, and in that way detect the bulge at the equator. That would be a physical fact requiring elucidation. We can even measure the rate of rotation with a Foucault pendulum or a gyroscope. (Foucault pendulum shown.) Hitherto this experiment had been regarded as a direct revelation of the rotation of the earth. If we interpret it in that sense, and if we accept Planck's criterion that an entity is physically real if it is measurable, then we find ourselves faced with the difficulty of reconciling the apparent physical reality of accelerated motions with the physical unreality of uniform translations. If you are standing in a tram when the brakes are suddenly applied, or when it rounds a corner quickly, the reality of non-uniform motion is brought home in vivid fashion to yourself and to the person on whom you fall; by no similar experience could we detect the motion of a tram running smoothly on a straight track. If the earth's atmosphere were opaque no experiment could reveal to us its translation through space, but the Foucault pendulum would reveal its rotation. It was from a sense of dissatisfaction with the philosophical aspects of this distinction that the general theory of relativity arose.

This sense of dissatisfaction forced Einstein to adopt the view that *all* motions are relative whether they be uniform or not; and he made this view plausible by the formulation of his principle of equivalence.

Imagine ourselves once more in our smoothly moving vehicle, and suppose that the brakes are suddenly and powerfully applied. Our common experience tells us that

we shall be shot forward out of our seats. The Newtonian view is that while the vehicle is moving we participate in its motion; during the sudden change in its state of motion our inertia carries us forward. The same result could have been attained in so far as our relations to the vehicle are concerned, if the state of motion of the vehicle remained unaltered and someone actually pushed us forward with a suitably adjusted pressure. We would speak of such an actual applied force as an impressed force as distinguished from the inertial force which we say is the cause of our upset. In Newtonian mechanics we keep these two sets of forces conceptually apart, our general attitude being that impressed forces are real in the sense that they originate in some external cause, while inertial forces are fictitious, but occasionally convenient methods of describing the effect of the motion upon us.

Einstein says that the Newtonian distinction between inertial and impressed forces is not philosophically valid. The effects produced are identical in the two cases; it is an unnecessary complication to ascribe identical effects to different causes.

Let us see how his ideas apply to questions of rotation. We shall consider in particular the rotation of the earth. If the earth were a perfect sphere and were not rotating, any body on its surface would, by virtue of gravitational attraction, experience a force pulling it towards the centre. That gravitational pull is an impressed force in the Newtonian sense, and we should then speak of it as the weight of the body. But the body participates in the rotation of the earth, and therefore describes a circular path, the radius of the circle being zero at the poles, and a maximum at the equator. Now we have seen that to force a body to move in a circular path we must provide a certain pull towards the centre of the circle, the magnitude of the pull depending upon the angular speed of the

rotation and upon the radius of the circle. The only possible source of that pull is the gravitational attraction of the earth upon the body. Part of such gravitational attraction is therefore consumed in compelling the path of the body to change from a straight line to a circle, and only the unconsumed portion appears as weight. It is clear therefore that the observed weight of such a body would be greatest at the poles and least at the equator.

Now we want to follow Einstein and regard the earth as being without rotation. If we take only gravitational force into account we shall not obtain the observed distribution of weight, for the weight of the body would then be the same at all parts of the earth's surface. We must imagine a set of centrifugal forces applied, the magnitudes of which vary from zero at the poles to a maximum at the equator; and the resultant of this axial centrifugal field superimposed upon the central gravitational field gives the actual observed distribution of weight. In the Newtonian sense the gravitational field is real, the centrifugal field fictitious. Einstein's principle of equivalence asserts that they must be considered as identical in every way, as we are unable to discriminate between them experimentally; all that we can observe is the difference between the two which we call weight. On this theory, therefore, the centrifugal field becomes one with the gravitational field, and a theory of gravitation based upon the relativity of all motions may not differentiate between them.

I must say that I do not share with the relativists their philosophical dissatisfaction with the apparent distinction between uniform and non-uniform motions. It is true that impressed forces and inertial forces are mathematically equivalent, for they are made equilibrate in the Newtonian equations of motion. But we must not base our philosophy upon a misinterpretation of the

abstractive methods of mathematical thought. The course of mathematical reasoning is something like this:—If certain relations were true, other relations would follow in consequence. Whether these relations be true or not, whether they be exemplified in Nature or not, does not concern us in the slightest; our field of discussion is the process of reasoning about relations. But when we become physicists and look around the external world we see certain relations variously exemplified; indeed it is this diversity of Nature that makes life interesting; only relativists and Trades Unionists hope to reduce everything to the same dull dead level of identity. We can, for example, describe all simple resonance phenomena in terms of a single differential equation expressing a fundamental relation between the variables concerned, but formal mathematical similarity does not entail physical identity. Would any of us care to assert that a mother rocking her cradle is the same thing as the reception of a wireless signal because they are covered by the same differential equation?

Within my own limited field of sense-perception I do discern a real difference between uniform and non-uniform motion; for all I have to do is to look out of my tram at the street and I shall at once see whether the cause of my losing my seat has been an impressed or an inertial force.

It is natural that we should endeavour to extend, rather than remove, the validity of this real distinction. But in doing so we find ourselves in trouble. We see that something is rotating, but whether it is the Earth or the Heavens we cannot say. We set up a Foucault pendulum, or a gyroscope, and again notice a rotation, but whether the instrument is rotating and the Earth at rest, or the Earth rotating and the instrument at rest we cannot say. Our windows looking into space yield no revelations, for

space is neither visible nor tangible like the street. All we can do therefore is to be grateful for the fact that inertial and impressed forces are mathematically equivalent, and in this sense we may accept Einstein's postulate; but it is an agnostic principle—a confession of ignorance; and it is hard to realize how a system of positive physical knowledge could be raised upon such a foundation.

SECOND LECTURE.

THE SPECIAL THEORY.

Section 1.—The Aether.

In the previous lecture we were concerned chiefly with a discussion of the theory of motion as formulated by a selected group of observers on the earth; but an adequate physical theory must account for other phenomena. We are constantly subject to influences originating in extra-terrestrial sources. Light, heat, gravitation are activities that are perpetually communicated to us across inter-stellar space. It is a characteristic feature of the human mind that it is seldom content with mere observation of such activities; it tries to conceive a suitable mechanism in terms of which it can describe such processes. To such a mind the idea of direct action between bodies at a distance through a void is repugnant and unsatisfying. Descartes could find no place for this idea in his system of philosophy; he assumed that force could not be transmitted otherwise than by direct impact or pressure. On his theory interstellar space must be a plenum, and not a vacuum; to the medium that is supposed to fill it is given the name "aether."

A record of the theories of the aether propounded at one time or another forms one of the most fascinating chapters in the history of physical science. We must, however, resist the temptation to linger by the way; we can only consider the position as it was about twenty years ago, just before Einstein startled the physical world.

Section 2.—Relative Motion of Earth and Aether.

We definitely know that light, heat and electro-magnetic disturbances can be transmitted from one body to another in space. We may either suppose that such transmission takes place in projectile fashion, or by wave

motion through a medium. Numerous experimental results have bestowed greater probability upon the wave-theory, and it is now generally accepted. It is difficult to imagine the process without some medium through which the waves are propagated, and such medium must have certain definite properties. It must have rigidity and it must have elasticity, else it cannot support wave-motion of the type contemplated. It must therefore have a certain degree of substantiality. If it is at all substantial we must consider how the earth is circumstanced with respect to it. Does the earth in its motion through space carry the aether with it, so that there is no relative motion between them? Or does it drag part of the aether along with it, as a ship in moving through water humps it up in front and leaves turbulent eddies in its wake? Or does the earth pass through the aether undisturbed, leaving no trace of its passage?

The first experimental result to indicate any choice between these alternatives was the discovery of astronomical aberration by Bradley, in 1728. Bradley noticed that the apparent position of a star varies slightly with the position of the earth in its orbit. The question puzzled him for two years, and then the explanation struck him that the phenomenon was due to the change in the relative motion of the star with respect to the earth, and to the finiteness of the velocity of light propagation. We can see the influence of the relative velocities very simply in this way:—Suppose a railway carriage is moving in the direction indicated, (Fig. 1), and that B is a bullet which has been shot from P in a direction perpendicular to the track. While the bullet is crossing the carriage, the carriage moves forward, so that B will not emerge at the point C but at a point D further back. The line

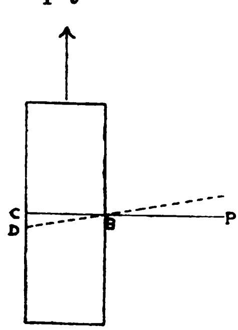

FIG. 1.

DB then gives the direction from which the bullet appears to have come as seen by an observer within the carriage, and that direction is not the same as for a man standing by the track. Now let T (Fig. 2) represent a telescope

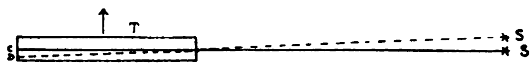

FIG. 2.

being carried round by the earth, and S a star. While light from S is travelling down the tube, the tube is carried forward by the motion of the earth, and the light will be brought to a focus, not at C the centre of the cross wires, but at D. To bring the image of the star on to the centre of the cross wires the telescope must be rotated slightly, so that it points in front of S instead of directly at S, and the star thus appears to be advanced from its true position. The phenomenon is purely one of relative motion. If the earth carried the aether round with it there would be no aberration, for there would then be no motion of the instrument relative to the ray of light. In 1845 Stokes suggested that the aether in the immediate vicinity of the earth is carried along by it, but that this effect diminishes as we recede from the earth, and that the aether becomes motionless at sufficiently large distances away. This theory gives the correct result for the aberration constant; but unfortunately the type of motion considered becomes possible only by assuming a degree of condensation of the aether in the neighbourhood of celestial bodies which is so enormous as to render the theory improbable.

The only hypothesis remaining is that of a stagnant aether, through which the earth freely passes, propounded by Young in 1804. This will not, however, account for all the known facts. It is, for instance, experimentally established that the velocity of light in transparent media such as glass or water is less than in

vacuo. Fresnel offered an explanation of this and other optical facts in terms of two further assumptions, viz., (1) a hypothesis already propounded by Young, that transparent bodies have the power of condensing the aether within them, so that the density of the aether within is increased, while the rigidity remains unaltered; and (2) that when transparent bodies move they carry with them only the amount of aether they have in excess over the normal density in vacuo; the result being that the moving body imparts to light passing through it only a fraction of its own velocity. The formula given by Fresnel was confirmed experimentally in 1851 by Fizeau, who passed a beam of light with and against a current of water, and found for the difference between the two velocities exactly the value deduced by Fresnel.

Hitherto no connection had been established between optical and electromagnetical phenomena, and the luminous vibrations in the aether were supposed to be of the same type as those encountered in the vibrations of elastic solids. Between 1860 and 1870 Clerk Maxwell developed a mathematical theory from which it appeared that electro-magnetic disturbances are propagated through the aether with a speed equal to a certain ratio involving fundamental electric and magnetic units. This ratio is, within the limits of experimental error, equal to the speed of light in vacuo. Maxwell boldly hazarded the opinion that light waves are electro-magnetic waves in the aether, and this brilliant synthesis unifies in one comprehensive scheme most known facts of optics and of electro-magnetics, enabling us to explain all radiational phenomena in terms of a single set of fundamental assumptions. The aether still stands as an essential part of Maxwell's theory—a stagnant aether through which the earth freely moves, but regarded as the seat of electro-magnetic, rather than of elastic stresses.

SECTION 3.—AETHER-CURRENT CANNOT BE DETECTED EXPERIMENTALLY.

We must now consider some of the endeavours that have been made to detect the motion of the earth through the aether. In terrestrial methods of measuring the velocity of light, a light flash is sent out from the observing station, and made retrace its path by means of a mirror. On a stagnant aether theory, in which the aether plays the part of a more or less substantial medium of transmission, it is not hard to see that the time for the double journey would depend upon the direction of travel of the light with respect to the aether current past the earth. The favourite analogy is that of a swimmer in a stream. Here, (Fig. 3), is a current of water running at say 3 m.p.h. as seen from the bank. A swimmer whose stroke can keep him going at 5 m.p.h. through the water is at O. He wants to reach a point A, 4 miles across the stream. For every 5 miles he goes through the water, the current carries him down 3; he must therefore aim at a point P, 3 miles upstream; he will then find that in an hour he will reach A. The same considerations apply to the return journey; and the swimmer takes in all 2 hours to go to A and back. Now let him swim 4 miles down-stream and back. He makes 5 m.p.h. through the water, the current carries him 3 m.p.h. down stream, so that he takes but half an hour to do his four miles. Coming back, however, it is not such easy going. For every 5 miles he moves through the water the stream carries him back 3; the return journey therefore takes him 2 hours. Taking the double journey in each case there is a difference of half an hour in favour of the cross-current swim.

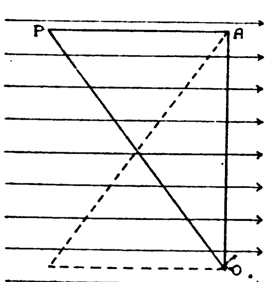

FIG 3.

We can express this difference more generally by an algebraical formula. From it we find that the slower the current, the smaller is the difference; if there is no current, there is no difference. The faster the current, the larger is the difference. When the speed of the current is the same as that of the swimmer the formula fails us altogether; the physical significance of this fact is that under such circumstance the swimmer could not go up or across stream, he cannot make headway against the faster current; in such case, therefore, the formula is not applicable.

In 1881 Michelson experimented with light beams as swimmers in the supposed aether current. From a point O (Fig. 4) he sent two beams along perpendicular arms, reflecting them by mirrors, M_1 and M_2, thus bringing them back to the starting point. If the earth were moving in the direction OM_1, then the aether current would be in the direction of the arrows. If the arms OM_1 and OM_2 were of equal length, then on the analogy of the swimmer that we have just considered, the time for the double journey OM_1 should have exceeded that for the double journey OM_2. Michelson was surprised to find that there was no observable difference in these times. The effect sought is extremely minute, for the difference between the times would be about the 100-millionth part of the time of either journey. No wonder Michelson's result was regarded with distrust. In 1887 Michelson and Morley repeated the experiment, refining the procedure to such an extent that they knew they could detect even one-fortieth part of the expected difference. Again the result was negative. Once more, in 1905, Morley and

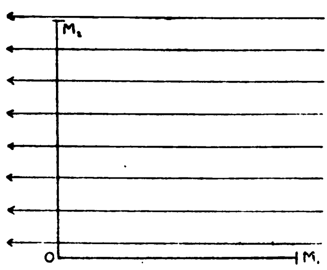

FIG. 4.

Miller repeated the experiment with additional refinements; but although their apparatus would have revealed even the hundredth part of the expected time difference, there was still no difference observable. The Michelson-Morley experiment, therefore, fails to reveal the slightest trace of an aether current past the earth. Confirmation, or refutation, of this has since been sought by other experimenters. Some of these results, such as those obtained by Lodge in 1892 and 1897 are in disagreement with the Michelson-Morley result, and point to a stagnant aether; others such as those of Trouton in 1902 and of Trouton and Noble in 1903 confirm the Michelson-Morley result and point to an absence of relative motion of the earth and the aether.

SECTION 4.—THE RECONCILIATION OF DISCREPANT RESULTS.

Here then we have three sets of experimental results forcing physical theory into an impossible position. Aberration and the Lodge experiments require a stagnant aether, giving rise to a relative motion of matter and aether as the earth moves in its orbit. The very sensitive experiments of Michelson and Morley and of Trouton and Noble fail to detect the slightest trace of such relative movement, and therefore require an aether moving along with the earth. Occupying an intermediate position we have the Fizeau and analogous experiments pointing to a partial drag of the aether by transparent moving matter.

Two attempts have been made to reconcile these divergent experimental results. We have on the one hand the positive *ad hoc* hypothesis propounded almost simultaneously by Fitzgerald of Dublin, and by Lorentz of Leiden, in 1892, to the effect that the motion of matter relative to the aether causes a definite change in the measured dimensions of the moving matter; and, on the other hand, there is the very general negative hypothesis

suggested in 1905 by Einstein that the aether, if it does exist, cannot be of the semi-substantial nature that we have ascribed to it, and that consequently motion of matter relative to the aether is meaningless, and to talk of an aether current is an absurdity.

To appreciate the significance of these suggestions we must return to the Michelson-Morley experiment.

SECTION 5.—THE FITZGERALD-LORENTZ CONTRACTION.

The Fitzgerald-Lorentz hypothesis is that all lengths contract in a certain ratio when turned from across to along the aether current; so that each arm of the apparatus becomes the shorter when placed along the current, and thus automatically counterbalances what would otherwise be a longer time of passage. We must be sure that we understand this contraction if we are to realise why it is that we cannot hope to measure it, although it does affect the light beam passing through the fixed aether. When a length is measured across the aether current the aether is not moving in that direction, so that it may be interpreted as a length measured when at rest in the aether. When a length is measured along the current it indicates a length in a system moving through the aether and measured from the fixed aether. If now we were to travel along with the apparatus and use a metre rod or other measuring instrument, as we lay it alongside the cross current arm it would be at rest in the aether, and the length thus measured would be expressed in terms of our fixed aether standard. But when we lay it alongside the other arm the measuring rod would also contract and in exactly the same ratio. On account of this compensating variation in the length of our measuring rod, no contraction would be observable by us no matter how the apparatus were moving. If, however, we were fixed in the aether, and could observe the apparatus moving past,

and if we measured the lengths with our fixed standards, then the length of the arm moving along the current would be less than that obtained if the apparatus were still. The light beam measures distances with respect to a fixed aether, hence it will detect the contraction. In this way an explanation may be given of the failure of the M-M experiment to detect relative motion between earth and aether.

This Fitzgerald-Lorentz contraction is, at first sight, a somewhat startling hypothesis, and many physicists question the possibility of its reality. To my mind there is no difficulty in accepting its reality. We are quite accustomed to changes in dimensions effected by physical causes. If you cool a rod it will contract; revolve it about one end, it will expand. Squeeze a steel sphere between your finger and thumb, it will no longer be spherical. Now the dimensions of bodies must ultimately be determined by the cohesive forces exerted between its particles, and these are of electrical nature. Electrical processes are indubitably modified by the passage of the bodies exhibiting them through the aether; it would not, therefore, be surprising to find that cohesive forces, and consequently, the dimensions of bodies were affected by such motion.

Attempts were made by Rayleigh, in 1902, and by Brace, in 1903, to verify by optical means the existence of this contraction; their results, however, were negative; no trace of double refraction could be detected in transparent substances moving through the aether. That null result need not disturb us; it means that the contraction is a natural process of readjustment within the constituents of the transparent body such that it remains optically isotropic, and is not a forced distortion such as would be produced by external pressure.

This contraction hypothesis supplemented Lorentz's

earlier electron theory. According to that theory the physical universe is constituted of two entities—the aether, and the ultimate constituents of matter which are electrically charged particles called electrons. The electrons are embedded in the aether, and react with each other not directly, but through the mediation of the aether. The aether is the aether of Maxwell, and is such that no part of it moves relatively to any other part. The theory gave a complete and satisfactory account of optical processes in stationary media, explaining in terms of these fundamental entities phenomena such as reflection, refraction, dispersion, radiation. When applied to moving media it gave the Fresnel formula, but the physical interpretation thereof is no longer an aether drag as suggested by Fresnel, for no aether can be carried along by the moving substance; it is now simply a change due to the motion of the electrons of the substance. The theory also showed that the inertia, or mass, of an electron, and therefore, of matter, is not constant, as assumed by Newton, but depends upon the velocity of the electron relative to the aether—a result of singular interest and importance since confirmed by experiment.

Section 6.—The Lorentz transformations.

Working always with the physicist's concept of fixed, absolute standards of distance and time intervals specified with reference to the stationary aether, Lorentz next set himself to examine how the laws of electro-magnetics would alter when referred to different inertial systems of reference. It was at once evident that they were not invariant for the Newtonian transformations. Lorentz discovered a set of relations connecting the aether scales and the moving scales, since then always known as the Lorentz transformations, for which the laws of electro-magnetics as determined for different inertial systems would remain invariant.

Suppose we have a rod fixed with us in the aether, let us measure it with our fixed standards, and say its length is found to be 2 metres. Now send it moving from us in the direction of its length; to us it would show a contraction depending upon its speed through the aether. An observer moving with it, however, would still report its length as 2 metres, for he would find that his metre scale, which had also contracted without his being conscious of it, still required exactly two applications to determine the length. His unit distance would appear too small, and his reported lengths would seem to us too large, as we can detect the contraction and he cannot; his readings would have to be divided by a certain quantity depending upon his speed in order to coincide with ours. Lorentz found also that he would have to modify the time scale applicable to the moving system in order to obtain the desired invariance. This new time scale he called the local time, but did not attach particular physical significance to it; it was a mathematical fiction designed to render the fundamental laws of electro-magnetics and optics identical for the stationary and the moving systems.

SECTION 7.—EINSTEIN'S POSTULATE OF CONSTANT LIGHT-VELOCITY.

It was at this stage that Einstein, in 1905, took part in the development of the theory. Many interpretations of Einstein's views have been published, some of them, I fear, difficult of comprehension; I give you what seems to me the most satisfactory way of looking at it.

Einstein began on the formal mathematical side. He noticed that the velocity of light is invariant with respect to the Lorentz transformations; that is to say, whether we measure the velocity of light in the fixed aether with the corresponding absolute distances and times, or in a

moving system with the local distances and local times attached to that system, the result is the same. Here he said, in effect, is a definite recognizable fact of nature, the propagation of light; let us take that as our basic phenomenon; let us say that its velocity relative to any observer is constant, and let us take that velocity as our unit velocity.

We have seen that with things as they are in the physical world, the fixation of a unit velocity is no more arbitrary than the fixation of unit length or of unit time, so that we need entertain no feelings of philosophical repugnance towards that aspect of Einstein's theory. The difficulty seems to me to be the physical one of accepting the constancy for all observers of the velocity proposed as unit. Many relativists boldly assert that constancy as an established experimental fact. Einstein himself[1] asserted it on the ground that otherwise different observers would obtain different results for that velocity; we saw however that such contingency need cause us no alarm, for it violates none of our previous physical beliefs. Schlick[2] says that "the M-M experiment teaches us that light propagates itself in reference to the earth with equal velocity in all directions." But the teaching is not so convincing as one would wish, for there is the alternative explanation of the Fitzgerald-Lorentz contraction, which accounts for the null result and yet allows the velocity of light to vary in different directions. Freundlich[3] says: "To all appearances this velocity plays the part of a universal constant in physical nature; every observer finds the same value for the velocity of light irrespective of his state of motion." But I should think that the probable error of the determinations of the

---

[1] See p. 19.

[2] *Schlick*: Space and Time. p. 9. Oxford (1920).

[3] *Freundlich*: Foundations of Einstein's Theory of Gravitation. p. 4. Cambridge (1920).

velocity of light is large enough to render that statement uncertain, and in any case experiments specially designed to detect that influence are rendered useless if we admit the Fitzgerald-Lorentz contraction.

I always think that the great beauty and harmony of Einstein's theory are marred by these attempts to found it upon a rigorous positive experimental basis. If we are to appreciate fully the theory without any reservations, we must simply postulate as a matter for belief, or otherwise, the constancy of the velocity of light for all observers; we cannot prove it experimentally, but we can say that no physical facts are known which cannot be reconciled with the postulate. We can justify it in this way.

In following out the consequences of Einstein's postulate we must eradicate from our minds all analogy with wave propagation through a more or less material medium. If we drop a stone into a still lake, a circular wave will spread out from the initial point. An observer, A, (Fig. 5), fixed at that point in the water will find the different parts of the wave equidistant from him at all subsequent times; the wave-velocity will be constant for him in all directions; call that velocity c. Suppose now that we have a second observer, B, who was beside A

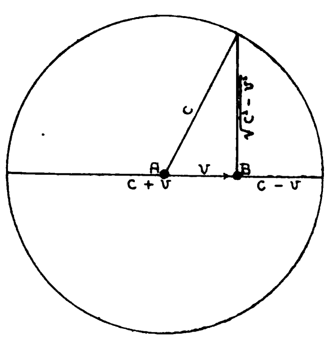

Fig. 5.

when the stone dropped, but is moving through the water away from A, with speed v. In one second from the dropping of the stone the part of the wave in front of B will be distant $c-v$ from him; that behind him, $c+v$; while the portions on his right and left will be distant $\sqrt{c^2-v^2}$. So that the velocity of the wave relative to B will be

different in different directions. That is the type of occurrence we should experience with a more or less material medium propagating luminous disturbances; if we believe in the existence of such a material medium, we cannot accept Einstein's postulate. But in the lake we could definitely state, and prove, that B was moving through the water; whereas we have seen that we cannot definitely state or prove that anything is moving through the aether.

Let A and B now represent two observers moving relatively in space. As neither can detect his motion relative to an aether, let us eliminate the aether entirely from our considerations. A and B will each state that he is at rest and that the other is moving, for only relative motion between the two is significant. Imagine now, that at a moment when A and B are together in space, a light flash is sent out from that point. Each thinking that he is stationary, will assert that he is remaining at the point from which the flash has been emitted, and that the other is moving away from that point. Had there been a third observer, C, there at the same instant he would have said that both A and B were wrong; that they were really moving away from the centre of the flash, while he, C, remained fixed at that point. These diverse descriptions of a single event are quite in accord with the known egocentric nature of our sense-perceptions. *If I imagine I am not moving in space; if I imagine that space is isotropic,* as I surely must in the absence of strong reason for believing the contrary, *and if a light flash is emitted from the point where I am, the only possible account I can give of it is that it moves away from me at constant speed in every direction.* This is, I think, the proper justification for the Einstein postulate; we now see it as another aspect of our fundamental agnostic principle, that we cannot detect motion through space.

From this point of view, Einstein's theory seems pre-eminently logical, and therefore consistent. Moreover, it gets us out of many of our previous difficulties which were due to the fact that our interpretations were expressed in terms of an aether current, whereas our experiments, although sufficiently sensitive, persistently refused to reveal that current. Since there is no aether current, and therefore no aether as we have hitherto imagined it, we can accept the Einstein postulate without imposing any strain upon our powers of acceptance; indeed, we can hardly do otherwise. The postulate at once reconciles the M-M experiment with other optical results. For the times taken to travel along the arms of the apparatus are constant no matter how they are turned, because the velocity of light is constant in all directions; no variation was revealed by the experiment because there can be no variation to detect.

SECTION 8.—RELATIVITY OF TIME- AND OF DISTANCE-INTERVALS.

How then can we reconcile the statements of our moving observers, A, B, and C that each remains at the centre of a light wave in space although they are moving relatively to one another? We must ascribe to each the same fundamental mental structure, so that if their positions were interchanged, the same descriptions would ensue. Either we must abandon all attempt at correlating their observations, in which case a physical theory would be impossible; or we must assume that their scales of measurement, both of distance and of time, differ one from the other, and that their variations of scale depend upon their relative motions. That is the new idea which we must assimilate, and for which our preliminary discussion has in part prepared us—that intervals of distance and intervals of time are not fixed invariable

things in themselves, but depend upon the system of reference to which they are referred.

So far we have fixed only our unit velocity—that of light. (Such unit would, in practice, be far too large, so we shall sub-divide it arbitrarily, and say that it contains c of our smaller units.) We may now fix either our unit distance or our unit time in any way we please. Imagine our unit time fixed by some suitable mechanism exhibiting regular periodicities: we may call it a clock. In a system, S_1, gather together at a point A_1, a number of these clocks, and keep them there under observation for a period sufficient to ensure the uniformity of their time-keeping. Then distribute them at different points of S_1, fixed relatively to A_1, and assume that they keep time in precisely the same way as long as they are fixed anywhere in S_1. From A_1 send out a light signal at, say, every fifth time interval. Then a clock so situated that it receives a flash half an interval after every fifth, would be distant $\frac{1}{2}c$ units from A_1; one receiving the flash at every sixth interval would be distant c from A_1, and so on. In this way we arrange a distance scale in S_1. Let the same process be carried out in a second system S_2.

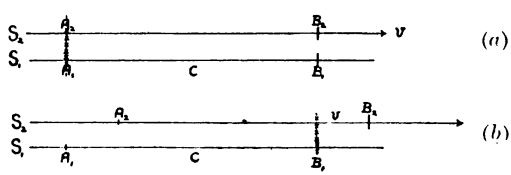

FIG. 6.

Now suppose S_1 and S_2 very close together, and moving with relative velocity, v. Let us attach ourselves first to the system S_1 (Fig. 6*a*). If it is zero time on the clocks at A_1 and A_2 when those two points are together, it will then be, for us, zero time at B_1 distant c from A_1, for the clocks in S_1 have been synchronized. The observer at B_1 distant c from A_1 can mark the point B_2 in S_2 opposite him at that instant. At the instant indicated in

the diagram let a light flash be emitted from the point A_1, A_2, and have mirrors placed at B_1 and B_2 so that it may be reflected back on reaching those points. We are looking on at the operation from S_1 and timing it by the units we have there adopted. In one unit of time the flash reaches B_1 and then starts on its return journey. In that unit B_2 will have moved forward a distance v, so that the position as seen from S_1 will then be as indicated in Fig. 6*b*. As we look from S_1 we see the point B_2 move ahead of the light flash, and the flash will appear to us to take longer than unit time to reach B_2. (In expressing it in this abbreviated fashion, I do not mean that we, individually, placed at A_1 will see these things. We have a whole line of observers in our system S_1, however, and they report whether B_2 or the flash reaches them first; so that when I speak of ourselves in S_1 I mean the co-ordinated results observed by all our stationary observers in that system.)

Consider now the reflected flashes as seen from S_1. The light reflected from B_1 will reach A_1 in unit time. The flash will have to travel further before it is reflected at B_2; then, as we look at S_2 from S_1 we shall see the point A_2 move towards the reflected flash. Hence the passage of the flash as judged from S_1 will appear to take longer than unit time to reach B_2 and less than unit time to reach A_2. To an observer in S_1, therefore, the times for the to-and-fro journeys between A_1 and B_1 will be identical, while those for the to-and-fro journeys between A_2 and B_2 will be different. There is nothing disconcerting in this discrepancy.

Let us now repeat the experiment, attaching ourselves to the system S_2. Since we are in S_2, that system is for us not moving, but S_1 is going backwards at a rate v. The to-and-fro journeys between A_2 and B_2 will now be judged by us to occupy equal times. But as we look from

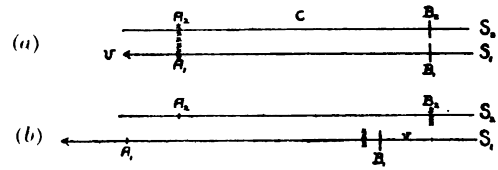

FIG. 7.

S_2 at S_1 we shall see B_1 coming to meet the light on the forward journey, and A_1 going away from the light on the return journey. Under our altered circumstances of observation the to-and-fro times in S_2 are equal, while the to-and-fro times in S_1 are unequal.

We conclude therefore that what are equal time-intervals in S_1 for an observer in S_1 are unequal for an observer in S_2; and that equal time intervals in S_2 for an observer in S_2 are unequal time-intervals for an observer in S_1. This then is the principle of the relativity of time—that the magnitude of a time-interval depends upon the motion relative to the observer of the system in which that time-interval occurs. As we see, it is an immediate consequence of the postulate of the constancy of the velocity of light for all observers.

The relativity of length follows at once. In the system S_1 we have a length to be measured. To do that we must have recourse to our fundamental unit, the velocity of light, and to our arbitrary unit, the clock interval of time. We send a light flash from one end of the rod to the other and time its passage by our synchronized clocks; that time multiplied by the velocity of light gives us the required distance. But an observer in S_2, looking on at this operation, would ascribe a different value to the time of passage across the rod; his value for the velocity of light is the same; hence his estimate of the distance, obtained by multiplying his time of passage by the velocity, would necessarily differ from ours. Distance measures therefore, as well as time measures, depend upon the relative motion of the observer and the thing to be measured.

I shall here introduce a term of which I shall make frequent use in what follows. I shall want to refer to this variation of scale in both distance- and time-intervals which takes place when the relative motion of the observer and the thing observed alters, and I shall speak of it simply as a change of *perspective*. This is a generalised application of the ordinary use of the word for it involves time as well as distance. As long as the relative motion in any direction is constant the corresponding perspective remains unaltered.

From considerations such as these, Einstein found the relations that must exist between distances and times in different systems, and obtained for them precisely the Lorentz transformations. That should not surprise us, for those transformations represent the conditions that must be fulfilled if the velocity of light is to be the same for all observers. There is however, this fundamental difference of interpretation : On the Lorentz theory the transformations represent a real contraction of matter when moving relatively to the aether, and a fictitious time applied to the moving system, in order to secure sameness of description; while on the Einstein theory, they represent actual differences of scale originating in the different circumstances of the systems.

Some of the difficulties of the so-called elementary expositions of the theory of relativity seem to me to originate in an unjustifiable mixture of these two interpretations. On the Lorentz theory there is absolute rest in the aether; on the Einstein theory there is none. On the Lorentz theory there is a real length and a real time; on the Einstein theory there are none. On the Lorentz theory there is a real contraction when matter moves through the aether; on the Einstein theory there can be no real contraction as there is no real length. On the Einstein theory no changes of structure take place but

changes of perspective occur, and the perspective in which we view things is determined by the motional relation to ourselves of the thing viewed. We may therefore describe things in terms of the one theory or of the other, but hardly in terms of a mixture of the two.

SECTION 9.—SIMULTANEITY, OBSERVED AND INFERRED.

I have endeavoured to explain these things to you—rather crudely, I fear, and certainly incompletely—explicitly in terms of measure relations, rather than in terms of the concept of simultaneity used by Einstein; because the concept of simultaneity is more involved, and leads more readily to misunderstanding than that of a relation between measure intervals. For example, in his pamphlet on "A Theory of Time and Space" Robb[4] wrote :—" . . . . Einstein made the suggestion that events might be simultaneous to one observer, but not simultaneous to another. This remarkable suggestion was at once seized upon, without it apparently being noticed that it struck at the very foundations of Logic. That a thing cannot both be and not be *at the same time* has long been accepted as one of the first principles of reasoning; but here it appeared for the first time in science to be definitely laid aside." And, again, in the preface to his recent book on "The Absolute Relations of Time and Space"[5] the same author says :—"From the first I felt that Einstein's standpoint and method of treatment were unsatisfactory. . . . . In particular, I felt strongly repelled by the idea that events could be simultaneous to one person and not simultaneous to another : which was one of Einstein's chief contentions."

The trouble arises from the fact that we are concerned with two kinds of simultaneity, namely, *observed*

---

[4] *A. A. Robb*: A Theory of Time and Space. p. 4. Cambridge (1913).
[5] Cambridge (1921).

simultaneity, and *inferred* simultaneity. Simultaneity of events can be observed only when they occur at one place; no simultaneity can be *observed* between events occurring at different places, it can only be *inferred* by means of some communication between the two places. The one is a primary basic fact instinctively recognized by the individual; the other is inference founded on some arbitrary measure scale. If I had, say, two clocks beside me, I could positively and definitely assert that they indicated zero time simultaneously, and my observation would be confirmed by any other observer who took my place. If however I move one of these clocks a thousand miles away, I can no longer observe whether identical readings are simultaneous or not. I must place another observer beside the second clock, get him to signal his readings to me, by light flashes, or by wireless, or some such means of communication, and then by calculation based upon the speed of our signals, and upon our distance apart, I could infer whether the readings were simultaneous or not.

Now, in the first case, where I have two clocks giving simultaneous identical readings beside me, it would, I admit, be absurd to have any other observer whatever in a position to assert that the clocks showed different readings to him at that instant. Had Einstein made such a contention as this, his theory would long ago have found well-merited oblivion. He fortunately does nothing of the sort. The Lorentz transformations show that *observed* simultaneities are simultaneous for all observers. If I were, say, to hold up my right and my left hands simultaneously, then, assuming that the fact could be of any interest to the universe at large, all observers would report simultaneity no matter what their circumstances might be. The Lorentz transformations likewise preserve *local* time-orders unaltered; that is to say, if two events, A and B, were to happen *here, where I am,* and I

observe that A precedes B, then for all observers however circumstanced, event A will precede event B.

It is otherwise, however, with inferred simultaneity. If I take a clock here, and a clock 1,000 miles away, I infer the simultaneity of their readings, founding my inference upon our distance apart, and upon the speed of our signals. Another observer, however, who is moving with respect to me, will infer that they are not simultaneous, for he founds his inference upon a different basis, his distance scale being different from mine.

Any difficulties concerning simultaneity that are encountered in the theory of relativity will be found to refer to such inferred simultaneities; but surely there can emerge no logical embarrassment from the fact that discrepant inferences are attained if the bases of inference are themselves discrepant.

SECTION 10.—CONSEQUENCES OF EINSTEIN'S THEORY.

The Einstein theory accounts at the same time for the aberration of light and for the null result of the M-M experiment; but it modifies the Newtonian Laws of Motion. This modification, however, is quite inappreciable for ordinary terrestrial velocities, and for such speeds the Lorentz transformations degenerate into the Newtonian transformations. But when we deal with velocities comparable with that of light the modifications become appreciable. If, for instance, we have a body B moving relatively to a body A with a speed of 200,000 km/sec, and a body C moving in the same direction relatively to B with a speed of 200,000 km/sec, then on the Newtonian theory we should say that the velocity of C relative to A would be 400,000 km/sec. On the Einstein theory that would not be correct; the velocity of C relative to A should be only 277,000 km/sec. The explanation is simple. The velocity of B with respect to A is expressed

in A's B-perspective; the velocity of C relative to B is expressed in B's C-perspective; and these velocities must be reduced to the same perspective before they can be added to give the velocity of C relative to A. We may go on adding as many subsequent relative velocities as we choose, and we shall find that, owing to the continual change of scale, we shall never find the total velocity relative to any body in the chain exceeding 300,000 km/sec, the velocity of light. I don't mind confessing that this circumstance as generally expressed puzzled me for years, and seemed to me a serious blemish upon an otherwise beautiful and compact theory. It is generally stated that you may add together any number of velocities you wish but the result cannot exceed the velocity of light. Eddington[6] puts it that this velocity "reminds us of the mathematicians' transfinite number Aleph; you can subtract any number you like from it, and it still remains the same." The statement is misleading. The operations of arithmetic are, fortunately, independent of our perspective; it would be obvious folly to assert that we might subtract any number we like from 300,000 and have as remainder 300,000. It is not a question of playing odd tricks with numbers; indeed, I think we may sometimes quite legitimately infer the existence of velocities greater than that of light. If we take the example I have just given you of the three bodies A, B, and C, and consider the velocities relative to B. The velocity of A relative to B is $-200{,}000$ (for it is a striking feature of the Lorentz transformations that the velocity of A with respect to B has the same numerical value as the velocity of B relative to A) and that of C relative to B is 200,000. B will infer that, in his units, the difference between the velocities of A and C will be

[6] *A. S. Eddington*: Space, Time and Gravitation. P. 59. Cambridge (1921).

400,000. We cannot directly add velocities referred to different systems of reference, because they mean different things; they are not measured on the same scale; they must be reduced to a common perspective before they can be added. Looked at from that point of view I find no difficulty, and no paradox. The proper mathematical analogy seems to me to be rather that of a convergent sequence, the convergence being conditioned by the diminution of the successive perspectives, and the limiting sum of the sequence being the velocity of light.

We must not allow a love of entertaining paradox to interfere with scientific precision of statement. Here is another quotation from Eddington[7] (he is dealing with the transformation from one system of reference to another): " It is a favourite device for bringing home the vast distances of the stars, to imagine a voyage through space with the velocity of light. The youthful adventurer steps on to his magic carpet loaded with provisions for a century. He reaches his journey's end, say Arcturus, a decrepit centenarian. That is wrong. It is quite true that the journey would last something like a hundred years by terrestrial chronology; but the adventurer would arrive at his destination no more aged than when he started, and he would not have had time to think of eating. So long as he travels with the speed of light he has immortality and eternal youth." It is a delightful idea that eternity might be interpreted as a Lorentz transformation; but unfortunately that transformation is not valid for relative velocity equal to that of light; it then entails division by zero, which is not a permissible operation. And even on the purely physical side it cannot hold, for reasons similar to those I mentioned in the case of the swimmer and the stream. The correlation of

[7] *op cit.* p. 26.

Section 8 fails, because the light flash as seen from S_1 would never reach the mirror in S_2. Our voyager could send no signal to us, and could receive none; his world would be limited to his magic carpet and its load of provisions; he would probably base his time scale on the rate of consumption of his edibles, for that, I imagine, would be the most interesting occurrence in Nature for him; but he would no more achieve immortality than we.

Einstein's law for compounding velocities gives us the Fresnel formula for the influence of the speed of moving transparent substance upon the velocity of light passing through it. His theory also gives us the correct expression for the shift of the spectral lines in the light reaching us from a star compared with the corresponding lines in light from a terrestrial source; and, finally, it gives the Lorentz expression for the apparent inertia of a moving electron. It accounts therefore for the known facts of mechanics and of optics in terms of one fundamental hypothesis regarding the method of correlating the time scales and the distance scales of moving systems. The principle underlying that hypothesis, however, makes us feel somewhat uncomfortable; it deprives our aether of all substantiality, and leaves us an elusive shadow that helps us not at all in conceiving the mechanism of transmission of energy between one body and another. Nevertheless there are in the theory no *ad hoc* assumptions, no inventions of particular and peculiar mechanisms to account for the observed facts of physics; from the basic method of correlation the laws of physics follow naturally and inevitably. It is an intellectual achievement that will stand for all time.

SECTION 11.—THE MINKOWSKI INVARIANT.

The next step in the development of the theory was taken by Minkowski in 1908. It also began on the formal

mathematical side, and I am afraid that it will be a matter of some difficulty to explain it to you in non-mathematical terms.

The Lorentz transformations gave us a correlation between the descriptions of the same event as reported by different observers; Minkowski went further and effected a synthesis; he showed us how we could form a mental picture of the universe as a whole, irrespective of individual perspectives.

To appreciate his point of view let us first consider an elementary case in which we deal with spatial relations only. Here in this room, time is the same for you as it is for me, so that it is not a factor for present consideration. If I hold up a coin we see it in different perspectives. Some of us will say that it is circular; some elliptical; some who see it edge on will say that it is a narrow rectangle. Our descriptions all differ, nevertheless we all have a definite and, I think, identical notion of the spatial relations of the various parts of that coin. There must therefore be something specifically characteristic of those relations which is independent of our point of view, or invariant with respect to our perspective; otherwise we could not obtain a common mental picture of it. We may map the spatial relations of the elements of the coin by marking recognizable points on it, and finding their co-ordinates in any frame of reference. The relative position of two points will then depend upon the differences between the three co-ordinates of these two points. Such differences are not themselves invariant; they will naturally depend upon the orientation of our coin in its frame of reference. If I hold the coin thus, and measure my x-co-ordinate straight in front of me, then the difference between the x's for the points at the top and bottom will be zero; they are both at arm's length from me; but if, keeping the coin still in the same position I

were to lie on my back on the floor, and still measure my x's straight in front of me, the difference between the x's for the same two points would then be equal to the diameter of the coin.

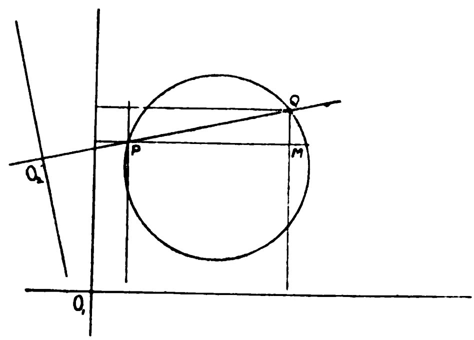

FIG. 8.

I can, perhaps, illustrate the procedure better for you in two dimensions. Here (Fig. 8) we have one aspect of the coin, with two points P and Q marked on its circumference. We refer it first to our axes of reference passing through O_1. The difference (dx) between the x's is PM; that between the y's is QM. Look at it now from another perspective, O_2, the difference between the x's is then the line PQ itself, while the difference between the y's is zero. No matter how the axes of reference are drawn we find that the length PQ is given by $\sqrt{(dx)^2 + (dy)^2}$. In three dimensional space the distance between two fixed points is

$$\sqrt{(dx)^2 + (dy)^2 + (dz)^2}.$$

That is invariant with respect to any position of our axes of reference; it is the invariant of Newtonian space; it represents a quantity specifically characteristic of the relative position of P and Q, for which the chance choice of axes is irrelevant. A knowledge of that invariant enables us to give a common account of the relative positions of all points on the coin no matter how we may be looking at it, and thus it effects a synthesis of our

different perspectives. Such syntheses yield the constructed space of physical science.

This example is very simple because we limit our problem to ourselves here, and thus our distance-scales and our time-scales are independent one of the other; time being the same for all of us it does not cause any variation in our distance measures. Such syntheses are purely local, but they suffice for most terrestrial purposes.

Imagine now that a man from Mars is looking down at us and sees us mapping out relative positions, not on our coin any longer, for that is on too small a scale, but, say, mapping out the positional relations between the highest point of the Town Hall, Johannesburg, and the highest point of the City Hall, Cape Town. He sees us lay off our co-ordinates, measure our dx, dy, dz, and calculate ds; he follows our example, but finds that his ds is different from ours, that it is not even constant. The local invariant which, for ourselves, serves the purpose of synthesizing perspectives by fixing position relations, will not suffice if we are to include the perspectives of moving observers. What would be for us an interval with extension in space only, is for the Martian observer an interval with extension in space and duration in time. Minkowski showed that, for the Lorentz transformations, the quantity

$$ds = \sqrt{(dx)^2 + (dy)^2 + (dz)^2 - c^2(dt)^2} \quad [c = \text{velocity of light}]$$

is invariant; and this is what we call the interval between two events. It embraces our own local Newtonian perspectives (for then $dt = o$, and $ds = \sqrt{dx^2 + dy^2 + dz^2}$ as before) and also those of all other observers who are moving in inertial systems with respect to us. It is something inherent in the relation between the events, independent of the framework to which we refer them; it is precisely the type of characteristic we require to

function as the basis of our synthesis. By taking $c=1$, in conformity with our previous notation, and by introducing a new kind of time such that $dt^2 = -dt'^2$, Minkowski showed that the above invariant takes the form $\sqrt{(dx)^2+(dy)^2+(dz)^2+(dt')^2}$. This is symmetrical in the four co-ordinates; the time interval dt' plays exactly the same part in it as the three distance intervals dx, dy, dz. Minkowski showed further that on this beautiful theory the Lorentz transformations are equivalent to a rotation of the axes of reference; "henceforth," he concluded, "space in itself, and time in itself, sink to mere shadows, and only a kind of union of the two preserves an independent existence."

On the mathematical side this modification is extremely elegant, but I don't think that I can quite follow Minkowski in his deductions from it. I can form no mental picture whatever of his new imaginary time. We use imaginaries every day in mathematics as symbols of direction, and the consequent theory is consistent as well as comprehensible. I can easily picture the result of turning a direction through a right angle, but I cannot ascribe any meaning to the rotation of time through a right angle. In my own mind, as I tried to explain to you last week, time-order and direction-order are fundamentally distinct conceptions, however much their measure relations may be interwoven, and in any synthesis I should naturally feel more satisfied were such distinction retained.[8]

[8] This could be effected by taking as invariant

$$ds = \sqrt{(dt)^2+(d\rho)^2}$$

where $\rho = ix+jy+kz$ = vector spatial distance between the events; but this is hardly a suitable place for discussing such a form.

THIRD LECTURE.

ON GENERAL RELATIVITY.

SECTION 1.—RESTRICTIONS OF THE SPECIAL PRINCIPLE.

In the previous lecture we discovered that transformations could be found for observers who are moving relatively to one another with uniform velocity in a straight line, such that the laws of optics and electromagnetics would be invariant for those observers. On account of this restriction it is called the *special* principle of relativity. In spite of the partial nature of its domain of validity it gave us as a first approximation Newton's laws of motion; and it accounted naturally for experimental results that had been puzzling physicists for years. In certain respects, however, the restricted nature of the principle is felt: it suffers from the same limitation as Newtonian mechanics inasmuch as it gives no definition of an inertial system other than in terms of motion with constant speed along a straight line; it takes no account of systems of reference which are accelerated, and, as we saw, rotating systems are of common occurrence; and it takes no account of gravitation, a universal property of matter. These restrictions Einstein removes in his general theory, according to which *all motions* are purely relative; and gravitation is brought into the scheme by means of the principle of equivalence, which asserts that there is no difference between impressed and inertial forces. Before we can discuss the general theory, however, we must say something about the nature of a manifold, and about the geometrical relations which may exist between the points of a manifold.

SECTION 2.—"INTERVAL" BETWEEN THE "POINTS" OF A MANIFOLD.

In all exact science something is measured. The measurements are referred to some conventional system of reference, and are expressed quantitatively by numbers which are multiples of some conventional measure-interval. We speak of all the possible states to be measured as forming a manifold, and the number of independent data required to specify any particular one of those states we call the order, or the dimensionality of the manifold. Thus, temperature is an entity that is the subject of exact measurement; any particular temperature is specified by a single number; various conventional scales of measurement (Fahrenheit, Centigrade) and various conventional starting points (ice and salt, melting ice) are used in practice, but no matter which convention is adopted, a single number referred to the appropriate system gives the temperature. Temperature is therefore singly-dimensional.

Pure tones, such as those emitted by tuning-forks, form a two dimensional manifold. Such tones differ in two respects, pitch and intensity. Each of these can be measured, and a knowledge of those two numbers fixes for us the character of the tone. (Ordinary musical notes such as we are accustomed to hear, constitute a manifold of more than two dimensions; for they differ not only in pitch and intensity, but also in quality; and quality is rather a complex attribute of sound which cannot be specified by a single number.)

The states of the entity defined by such sets of numbers we speak of as points in the manifold, and the sets of numbers which fix the points we call the co-ordinates of the points. We shall specify such co-ordinates by the letters x_1, x_2, etc.

From our definition it is clear that in itself and of itself

a manifold has no structure, no form; it is necessarily amorphous. Structure is conferred upon it only by the discovery of relations between its constituent parts; structure is thus a matter of experience. In the manifold of pure sounds which I mentioned just now, experience yields no knowledge of any fixed relation between two points; we can find no expression that enables us to compare the interval between two sounds, with that between any other two sounds. (I use the word interval here in the mathematical sense, involving both pitch and intensity, and not in the musical sense. Musical intervals are dependent upon pitch only; no account is taken of intensity. So that musically tones form a singly dimensional manifold, and the interval between two points of it has then a meaning, for it is simply the difference between the two values of the single co-ordinate—the pitch.)

The local space of our sense-perceptions is a three dimensional manifold, for three numbers are required to fix the position of a point in it. Experience yields us a basis of comparison of intervals between points in our local space, for we know from experience that rigid bodies can be moved about in our local space without undergoing any change of shape. The length of a rigid rod is then a definite thing no matter how it may be oriented with respect to our axes of reference; or, what comes to the same thing, the distance between two fixed points is constant no matter how the axes are chosen with respect to them.

When we enquire how it comes about that some manifolds, such as our local space, have structure conferred on them by experience, and others, such as our sound manifold, have not, we find the answer in the circumstance that some manifolds are homogeneous and others are not. The local space manifold is homogeneous, because the co-ordinates which fix a point in it all mean

the same kind of thing, namely, distances measured along specified directions. The pure sound manifold is, on the other hand, heterogeneous, because the two co-ordinates are there distinct non-comparable quantities of different essential nature; we can find no structural significance common to pitch and intensity.

We have already seen that our world—the world of events—is a four dimensional manifold. An event is fixed in it by means of four co-ordinates; three of these are required to fix its "where," and the other fixes its "when." From the ordinary philosophical and physical point of view the manifold is heterogeneous, the "where" is essentially distinct from the "when"; no comparison is possible or conceivable between them. From this point of view it has no structure, and no meaning can be assigned to the "interval" between two points in it. But the special theory of relativity showed us that intervals of distance and intervals of time are not the independent things hitherto supposed. When all observers are taken into account distance involves time, and time involves distance; the co-ordinates become entities of the same nature, and the space-time manifold is consequently rendered homogeneous. The interval between two points in it acquires significance.

Section 3.—Euclidean and non-Euclidean relations.

To appreciate the geometrical notions involved in the treatment of the space-time manifold in the general theory of relativity, we shall find it convenient to return to the simpler case of two dimensions. We may specify the position of a point in a plane by means of two co-ordinates representing the distances, x_1 and x_2, of the points from two rectangular axes. The interval between two points is then given by $ds^2 = dx_1^2 + dx_2^2$, and that interval is then invariant for all such axes. If we referred

the position of a point of this two dimensional manifold to two axes which are not rectangular, we should find for the interval between two neighbouring points $ds^2 = dx_1^2 + dx_2^2 + 2k.dx_1.dx_2$, where the k depends upon the angle between the axes. We may also obtain a system of reference by dividing the plane into sections by means of concentric circles, and specify the position of a point by means of two co-ordinates one of which gives the number of the circle passing through it, and the other the angular distance one must proceed round the circle before reaching the point; in that case we should find for the interval between two neighbouring points $ds^2 = dx_1^2 + x_1^2.dx_2^2$. In each system of reference the interval is invariant, but the form in which it is expressed depends upon the nature of the system of reference.

We may combine the different systems of reference and write generally for the interval between two neighbouring points of any two-dimensional manifold.

$$ds^2 = g_{11}\, dx_1^2 + g_{22}\, dx_2^2 + 2g_{12}\, dx_1.dx_2$$

where the values to be ascribed to the g's depend upon the system of reference to which the x's refer. In the case of rectangular axes, for example, $g_{11} = g_{22} = 1$, $g_{12} = 0$; with the oblique axes $g_{11} = g_{22} = 1$, $g_{12} = k$; while with the circular system of reference $g_{11} = 1$, $g_{22} = x_1^2$, and $g_{12} = 0$.

We shall call the first of these cases, which is nothing other than the theorem of Pythagoras, that the square on the largest side of a right angled triangle is equal to the sum of the squares of the other two sides, the standard euclidean form. If we know the transformation required to pass from one system of reference to another we could find from the standard form the new values of the g's for the second system; or we can reverse the process and use the transformation to bring the g's of any given case back to standard euclidean form.

Sometimes, however, we may find values of the g's which refuse to be transformed back in this fashion. If the two points were points upon a spherical surface, and the two co-ordinates were the respective latitudes and longitudes of the points, we should find for the interval $ds^2 = dx_1^2 + \cos^2 x_1 . dx_2^2$, and no relation between the x's on the sphere and the x's on a plane can be found which will transform this back to standard form. The reason is that there is an essential difference between the metrical relations obtaining on the surface of a sphere and those obtaining on a euclidean plane.

The difference is this: on a plane surface there is no variation of metrical relations when the scale upon which the measurements are made varies; upon a spherical surface there is. To appreciate this fact let us take our distance ordering intuition as giving us a straight line on the surface within our own immediate field of experience. To find the straight line joining two distant points on the surface let us build up a path between these two points out of such elementary straight lines. It is clear that we may move along the surface from the one point to the other by an infinite variety of such composite paths; but amongst all possible routes there will be one which makes the sum of the elementary distances a minimum. That minimum route is the shortest distance between the two points as measured on the surface. If we had no knowledge whatever of points lying off the surface we should take that minimum route as our straight line joining the two distant points. For the present imagine that our knowledge is so restricted. Within our limited fields of experience we make measurements and deduce measure relations. We find, in particular, that for triangles drawn in our immediate neighbourhood, the sum of the angles is always 180°. If we draw various circles and measure them we find that, as far as we can tell, the ratio

of the circumference to the diameter is constant. The geometry of our immediate field of expérience is euclidean. No matter what the surface was upon which we carried out our measurements we should find that by taking a sufficiently small portion of it, the measure relations we deduced would be of this type. But when we try to piece together these limited fields of experience it is not hard to see that the large-scale results we obtain may well lead to different relations. If the surface were flat we should find no difference between large- and small-scale geometries. But suppose the surface were spherical, say, instead of flat; then we should find, on taking large triangles bounded by our supposedly straight lines, that the angles no longer added up to 180°. On drawing large circles on our surface we should find that the ratio of the circumference to the diameter diminished as the diameter of the circle increased. We should also find that there existed a circle of maximum diameter. Measure relations under such conditions, upon such a surface, would depend upon the scale of operations; large scale relations would differ from small-scale relations; large scale geometry would be non-euclidean. There is thus revealed an essential difference between a plane surface and the surface of a sphere.

Gauss showed that if one surface could be applied to another by bending without stretching, then the geometries on the two surfaces would be identical. Thus, a cylindrical surface can be cut and laid flat, hence for co-ordinate systems of reference upon a cylindrical surface transformations can be found to reduce the expression for the interval between two points to standard euclidean form. But a spherical surface cannot be laid flat however cut, hence on such a surface the interval will involve coefficients which cannot be so transformed, except for small regions of the surface.

The distinction is exemplified in the theory of cartography. As you are no doubt aware there are various methods of representing on a plane the metrical relations between points on the earth's surface; but none of them can represent all the metrical properties of the whole earth at the same time. Some of the systems are what we call conformal, that is to say they preserve relative positions unaltered over small regions; but the magnification changes from point to point, so that although within limited regions relative distances are unaltered, yet over larger regions there is necessarily distortion. That is a good illustration of the distinction that I wish you to realise. It is reflected mathematically in the behaviour of the g coefficients in the expression for the interval between two points on the surface; if those g's can by mathematical processes be transformed into the standard euclidean form, then the surface we are dealing with is either a euclidean plane, or is applicable thereto; if the surface is such that on it large-scale relations are non-euclidean, that fact will be reflected in the non-transformability of the g's.

Because our spatial perceptions are three-dimensional we can interpret such non-euclidean relations upon a surface in a very simple way; for, thinking in terms of three dimensions, we can easily realize how it comes about that the nature of large-scale relations depends upon the surface upon which the measurements are made. To us the "straight" lines upon the surface are not straight; they are shortest distances, or geodetic lines upon a surface that is curved in our space.

If we were to make that curvature diminish, the surface would become flatter. We might conceivably have a surface of such small curvature that it would be quite impossible for two-dimensional beings upon its surface to detect any variation in their geometrical relations; to

them it would be a euclidean plane, and yet it would be for us a curved surface in our three dimensional space, although the curvature would be small.

We can also interpret non-euclidean relations in another way. Let us consider the example given by Einstein of measurements made upon the surface of a rotating disc. Place an observer O_1 at the centre of the disc, and another observer O_2 at a point on the circumference. There is no relative motion between them in the direction of their join; hence no matter where we place O_2 on the circumference, O_1 and O_2 will obtain the same value for the length of their line of join. They will both describe the disc as a circle, and will ascribe the same value to its radius. The observers *are* moving relatively to each other in any other direction; their perspectives in any other direction will differ; consequently their determinations of the length of the circumference of the disc will differ. The ratio of circumference to diameter would then be no longer constant for all observers; it would depend upon how far away the observer was from the centre of the disc, and how fast the disc was rotating. Geometries based upon such measure relations would be non-euclidean, and their non-euclidean nature would originate in the different perspectives appertaining to the various points on it.

These concepts are extended, by analogy, to worlds of more than two dimensions. When non-euclidean relations are encountered in the four-dimensional world of events, Einstein interprets them in terms of a curvature of that manifold in a world of five dimensions. Of such curvature I can form no picture whatever; I cannot imagine the nature of the fifth dimension. The analogy is useful in so far as it gives us convenient abbreviated methods of characterising the intrinsic nature of different regions of the space-time manifold; but it otherwise

conveys no meaning to my mind. I shall therefore adopt the second method of interpreting non-euclidean relations in terms of perspectives, as it is quite comprehensible, and moreover, it follows naturally from our interpretation of the special theory.

SECTION 4.—THE INTERVAL BETWEEN POINT-EVENTS IN THE SPACE-TIME MANIFOLD.

Since it requires four co-ordinates to fix a point in the space-time manifold, we must expect the expression for the interval between two such points to be more complicated than in the simple two-dimensional case we have just considered. I write it down :—

$$\begin{aligned} ds^2 = {} & g_{11}\,dx_1^2 + g_{22}\,dx_2^2 + g_{33}\,dx_3^2 + g_{44}\,dx_4^2 \\ & + 2g_{12}\,dx_1\,dx_2 + 2g_{13}\,dx_1\,dx_3 + 2g_{14}\,dx_1\,dx_4 \\ & + 2g_{23}\,dx_2\,dx_3 + 2g_{24}\,dx_2\,dx_4 + 2g_{34}\,dx_3\,dx_4 \end{aligned}$$

but, fortunately, to understand the general trend of the argument it is not essential for you to grasp its full significance. What I want you to bear in mind is that (i) the four x's are the four co-ordinates, or numbers, which fix the position of an event in the space-time manifold, and that the four dx's are the changes that these numbers undergo in the passage from one event to a neighbouring one. The actual values of the x's and of the dx's will, of course, depend upon the system of reference adopted in the manifold. (ii) The 10 g's are coefficients which depend upon (a) the particular system of reference adopted, and (b) the intrinsic nature of the manifold itself. Insofar as (a) is concerned, the sets of g's are mutually convertible by suitable transformations, but differences which originate in the intrinsic nature of the manifold itself are not so transformable except for small limited regions at a time.

Now the special theory of Einstein, as interpreted by Minkowski, taught us that the interval between two

events in Nature as seen by different observers who are moving relatively to one another with uniform velocities could be written $ds^2 = dx_1^2 + dx_2^2 + dx_3^2 + dx_4^2$ (written in the previous lecture in terms of x, y, z and t'). As long as we remain attached to one such system and refer the events to that system we view them in the same perspective; we use a constant set of units; three of the x's are our space co-ordinates and one is our time co-ordinate. When we attach ourselves to another uniformly moving system we adopt a new perspective; our x's have still the same significance but they refer to different scales. The expression for the interval between the two events remains the same in spite of the change of perspective; and on account of the form of that invariant we say that this space-time manifold of the special theory of relativity is euclidean.

But in any general theory, systems which are not moving with uniform velocities must surely be taken into account; in such systems the velocity is changing; in such, therefore, the perspective changes, and the measure scales vary from place to place and from time to time.

Section 5.—The Minkowski world-line.

In the absence of the necessary mathematical equipment we can best follow Einstein's argument in terms of the Minkowski "world-lines." We shall simplify matters, initially at any rate, by considering a one-dimensional or line world. Movements are taking place along this world. If we watched these movements their tracks in their world would overlap, and, as a record of the happenings in that world, would be valueless. You know how this disadvantage is overcome in recording instruments such as the barograph. There the events along the line make a mark upon a moving strip of paper. That moving paper represents to us visibly the flow of

time. It is clearly a matter of indifference whether the paper moves past the recording point or the point moves over the paper. We shall consider our line world as moving past time, instead of time moving past our world. Everything in the line-world may be regarded as leaving a track of its passage through time; such tracks Minkowski terms *world-lines*. We may fix our ideas a little more precisely by imagining a timekeeper of some kind attached to everything so that its beats are marked automatically at equal intervals along the corresponding world-line.

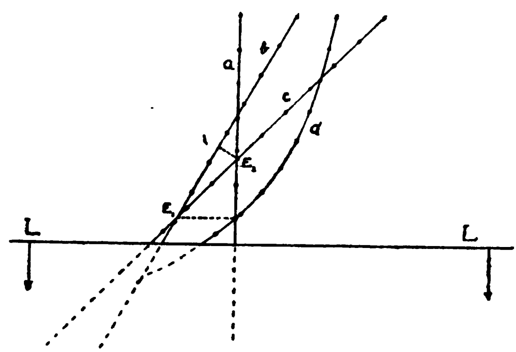

Fig. 9.

We from our superior space can watch the whole process. Here say (Fig. 9) is the world L moving through time as shown. A world line such as (*a*) would mean, *for us*, a stationary event in L. A line such as (*b*) would mean, for us, passage along L from right to left. A line such as (*c*) would mean, for us, passage in the same direction but at a higher speed. And we who look on from the outside can say that if (*a*), (*b*), (*c*) represent observers in L, (*a*) is stationary, and that therefore his line represents the true time scale.

But according to our fundamental hypothesis, (*a*), (*b*), and (*c*) will be quite unconscious of their own movements in L, and will have no means of detecting it. Each will say that he is stationary and that his world line

represents his passage through time, that the intervals along it are, for him, true time units.

Now consider the points marked E_1 and E_2. The former is the intersection of the world-lines of (*b*) and (*c*); it therefore represents the event which is the coincidence of (*b*) and (*c*) in space and time. Similarly E_2 represents the coincidence of (*c*) and (*a*). The interval between these two events would be for (*c*) a pure time interval, for he considers himself stationary, and sees (*b*) and (*a*) pass him. Geometrically the two events lie on his world-line; this represents a pure time-interval; in our example it is about $2\frac{1}{2}$ units. But for (*a*) the interval between the same two events is not a pure time interval. E_2 is on his world-line, but E_1 is not. He determines the time interval by measuring it on his own time-scale; that is, by dropping a perpendicular from E_1 on to his world line. The interval between the two events is then split up into a time interval of less than 2 units, and a distance interval given by the perpendicular distance from E_1 to his world-line. Observer (*b*) would divide the interval differently. Event E_1 lies on his world line; E_2 does not. Measuring time, as he must, along his own time scale he would get different values for his time and distance intervals.

This is not by any means a correct picture of the Minkowski method, for in this illustration I am using true time, *i.e.*, measured time, instead of Minkowski time; but you can see from it in a general way how coincidences in space and time, which form the subject matter of experimental science, are given by the intersections of world-lines, and how the relations between time-units and distance-units depend upon the inclination of the world-lines, that is to say, upon the relative velocity of the observers.

These world-lines are all straight, their straightness being conditioned by the uniformity of the corresponding

velocities in L. Things moving along L with non-uniform velocities will have world-lines which, as seen by superobversers like ourselves outside their manifold, will not be straight. A world-line like (*d*) we should interpret as motion along L from right to left with increasing speed. Now according to the relativity principle an observer moving like that, or in any other way, will be unconscious of his motion; if he can think in terms of world-lines at all, he will consider that his own world-line is straight; but that he is in a field of force; just as our passenger in the erratically moving tram, who is not allowed to look out at the street, may say that his vehicle is at rest, but that he is acted upon by forces arising from some unknown cause. As this observer's idea of straightness is derived from his own world-line he will say that other world-lines which we, as superobservers, consider straight, are to him curved. The best analogy I can think of is the ethical one. We all form our own ethical perspectives. Most of us will usually assert that our track through our ethical space does not deviate appreciably from rectitude; and yet we have seldom much hesitation in attributing marked crookedness to our neighbour's line. When we reflect upon it, however, we are obliged to confess that although his track does not seem straight in our perspective, it need not necessarily diverge from rectitude in his.

The ideas remain essentially the same, but the geometry becomes more complicated, when we deal with world-lines in manifolds of more than two dimensions. In such higher manifolds world-lines need not necessarily, and in general will not, intersect; when they do intersect, such intersections determine space-time coincidences, that is to say, physical events.

SECTION 6.—THE GENERAL RELATIVITY PRINCIPLE.

Our space-time manifold is given by the passage of our three-dimensional spatial world through time. The ideas and terminology remain the same, but I can draw no picture of it; for to obtain a picture of it would require space of four dimensions. What we can usefully do is to consider various sections of the manifold. If we keep time constant we obtain a section which is our instantaneous space; that would be a section perpendicular to our world-line. But we have seen that each observer will consider his own world-line to be straight, hence what are perpendicular sections for some observers will not be perpendicular sections for others. My instantaneous space will therefore be different from that of an observer who is moving relatively to me; space accordingly becomes a sectional aspect of a perspective in the manifold, and is robbed of all vestige of independent existence.

In the manifold as a whole there must exist all sorts and conditions of world-lines, giving the history of the universe at all times. If an observer, O_1, drew a picture of these world-lines as he imagines them, that picture would differ from the picture drawn by another observer, O_2, who is moving relatively to O_1, for they view things from different perspectives. If their relative velocity were constant the distortion of one picture with respect to the other would not be a very serious matter, for straight world-lines in the one picture would still be straight world-lines in the other although with different inclinations. If, however, their relative velocity is not constant the distortion becomes serious, for what are straight lines in the one picture become curved lines in the other; and the more the relative velocity of the two observers changes the more serious becomes the divergence between the two pictures. Now as we wish to try to synthesize the accounts of all possible observers, we must

allow for all possible distortions of the pictures of all world-lines in the manifold. At first sight it seems a pretty hopeless thing to attempt. The distortion depends upon the motion of the observer, and we are to permit the existence of all possible observers; we must therefore admit all possible distortions, and there seems to be nothing definite left for us to hold on to. There is something, however.

All objective events are of the nature of coincidences in space and time, and no change of perspective can possibly alter the fact of such coincidence. Suppose someone were to give me a sovereign; that would be an event. My world-line and that of the sovereign would exhibit a coincidence in space and time, not only as seen by myself, but also as seen by less fortunate observers. They would describe our world-lines in many different ways according to their perspectives, but no variation of perspective could alter or obscure the fact that such coincidence had taken place. No matter how the world lines in our world pictures may be drawn, they must all exhibit the same coincidences. Now the laws of physics are expressed in terms of such space-time coincidences, and these are independent of the perspective from which they are viewed; hence the laws of physics must be so formulated that they are invariant for all observers.

That is one method of formulating the general principle of relativity, namely, that the measure system of any perspective must be so chosen that the laws of physics are the same for all perspectives.

SECTION 6.—EINSTEIN'S LAW OF GRAVITATION.

This general principle, however, does not allow of the formulation of any law at all; we must start with some perspective as known. Einstein therefore assumes that (i) there must be regions of the universe in which the

principle of inertia holds, otherwise we should hardly find deductions based upon it so consistently verified in astronomical observations; we may therefore postulate that inertial systems exist somewhere; (ii) we do find that the special principle of relativity is valid within suitably chosen domains, which may be infinitesimal with respect to the universe at large, but are still finite with respect to our perceptions. As we saw, the special principle accounts for the facts of mechanics and of optics in a manner which justifies our acceptance of it as a law of nature, in the sense in which we have used that term. Within such regions of the space-time manifold, choosing co-ordinate systems of reference as indicated by Minkowski, we can find observers who would describe the interval between two neighbouring events as $ds^2 = dx_1^2 + dx_2^2 + dx_3^2 + dx_4^2$; and we say that for these observers such regions are euclidean with respect to their measure relations. We shall call such conditions fundamental, because they are the simplest and most symmetrical that we can conceive. We can now evaluate the interval between two distant points of the manifold by moving continuously from the first to the second, adding together all the elementary intervals that we get in moving from one point to a neighbouring one. One of all these possible paths will be a minimum route, a geodesic in the manifold. Such geodesics are the natural tracks of events within the manifold; they are the world-lines of particles removed from the influence of all other matter. As that is about the most elementary dynamical conception we can form, we call such tracks within the euclidean domain "straight lines."

The mathematical processes whereby these considerations are extended are somewhat complicated; I can only endeavour to give you a slight idea of their nature by means of some analogy, although argument by analogy is seldom convincing and always dangerous.

I have already explained to you how we construct our space by means of the invariance of the distance between two points in it. Suppose now we were to look upon a well-designed building; we are conscious of harmonies of proportion, of symmetries of form, of grace of outline, and other characteristics which are our aesthetic interpretations of certain measure relations existing in our constructed space between the different parts of the structure. Let us now regard the same building through a window whose glass is of the quality generally used by Johannesburg builders. Our sense of proportion finds the result offensive. Lines that were straight in our former perspective are now wavy; distances that were equal are now unequal; symmetries become distorted. Still if the glass were not too bad, we should find that to each point of the building as seen in one perspective there corresponded a point in the other. By examining the measure relations between various points in the two perspectives we could establish the transformation which converts the one picture into the other. If we look through many such panes of glass we can get all sorts of distorted images with their varying laws of transformation correlating a specified point in one image with the corresponding point in another. Even though we were unaware of the presence of the anisotropic glass and consequently of the cause of the distortion we should naturally take our first picture as fundamental, because it satisfies most completely our sense of proportion, and we should derive the other pictures from the fundamental one by the appropriate transformations.

I postulated that the glass of our window panes should not be too bad; generally it is. We often find a small region of the pane such that we can see nothing at all through it; no matter how we looked at the building through that glass there would be a patch of the picture

missing. By moving our position and thus viewing it from different perspectives, we can move that missing patch from one part of the building to another, but we can never transform it away completely. Were we ignorant of the existence of the glass and of its peculiar properties we should say that there was something in our perspective whose existence was revealed to us by its refusal to be removed entirely from our vision.

When there is no such patch we can get point-to-point correspondence between the two appearances for the whole building; when there is such a patch we can get point-to-point correspondence for any portion we wish, by so adjusting our perspective that the patch is off the region in question; but we cannot then get point-to-point correspondence for the whole building; the patch is always somewhere on it.

You must not push this analogy too far; it is, I admit, somewhat far-fetched. But if you bear it in mind you may be able to grasp an important distinction that must be made between two types of transformation.

We shall now consider a second observer in the euclidean region of the manifold, whose motion with respect to the fundamental observer is not uniform. He will be unconscious of his motion; he will recognize no distortion of his perspective; he will deem himself to be at rest and his own world-line to be straight. He will consider curved the straight line determined by the fundamental observer; but that line will still be for him the path of a particle removed from the influence of all other matter; it will therefore be a geodetic line in his perspective, for the geodesic is something that depends only on the intrinsic nature of the manifold and not upon the perspective adopted, and we are considering in the meantime only a euclidean region of the manifold. To the interval between two neighbouring points on this

geodesic he will ascribe a value $ds^2 = g_{11}\, dx_1{}^2 + \text{etc.}$, where the g's will depend upon his perspective. Different observers within this region will use different perspectives, and therefore different coefficients in their expression for the interval ds. We may accordingly take these coefficients as defining the perspective; any one set of g's must be transformable into any other by the appropriate change of perspective. One cannot tell simply by looking at them whether these sets of g's satisfy the condition of being mutually transformable or not; if they are, they have to satisfy certain rather complicated mathematical relations. I shall want to refer to this all-important fact that perspectives within a euclidean region are mutually transformable, while perspectives outside such domains are of a different nature. The relations conditioning such transformability can be most succinctly expressed in terms of the vanishing of certain quantities which are called Riemann symbols. For a four-dimensional manifold there are 20 such symbols. If these all vanish for a certain perspective then it belongs to the euclidean domain; if they do not all vanish for that perspective then it belongs to some other region of the manifold.

Let us now consider a region of the euclidean domain and imagine a number of isolated particles moving therein. An observer attached to one of those particles would say that their world-lines were straight; an observer whose system of reference was rotating would say that their world-lines were not straight; he would attribute their distortion to a field of force; and on removing that inertial field by a suitable transformation he would find that not one but all of the tracks would be restored to their proper rectitude.

Consider on the other hand a number of isolated particles moving in what we would call a real gravitational

field. Say S (Fig. 10) is a focus of gravitational attraction. Particles under the influence of S experience forces whose directions change from place to place. An observer fixed relatively to S at O will say that the world-lines of those particles are all curved. If he moves

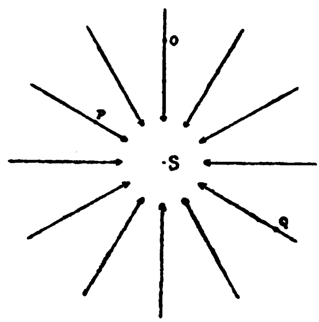

Fig. 10.

his system of reference with the acceleration characteristic of the field at O, then the world-lines of the particles in his immediate neighbourhood will be made straight, for they will then move uniformly with respect to him. But for particles at other points of the field, say at P and Q, the world-lines will not be straight; for obviously the compensation that removes the field at O cannot at the same time compensate the fields at P and Q, which are in different directions. An inertial field therefore can be removed completely and entirely by suitable transformation, whereas a gravitational field can only be removed from a particular neighbourhood to reappear elsewhere.

One would have thought that this is a real distinction which is confirmed by mathematical treatment, for the Riemann symbols vanish for the whole of an inertial field, whereas they do not vanish for the whole of a gravitational field. But an extraordinary feature of the Einstein theory is that he and his followers sweep it away as being philosophically intolerable. Eddington[1] puts it in this way :—" It may be urged that we have not stated

[1] Eddington: *op. cit.* p. 67.

the case quite fairly. When we choose the non-rotating observer, the centrifugal force disappears completely and everywhere. When we choose the occupant of the falling projectile.... so far from getting rid of the field of force, he has merely removed it from his own surroundings, and piled it up elsewhere. Thus gravitation is removable locally, but centrifugal force can be removed everywhere. The fallacy of this argument is that it speaks as though gravitation and centrifugal force were distinguishable experimentally. It presupposes the distinction that we are challenging. Looking simply at the resultant of gravitational and centrifugal force, which is all that can be observed, neither observer can get rid of the resultant force at all parts of space."

The philosophical basis of this criticism is put more explicitly by Freundlich,[2] who regards it as a fundamental postulate in the mathematical formulation of physical laws. He says that "in the formulation of physical laws only those things are to be regarded as being in causal connection which are capable of being actually observed." He proceeds to say that the reason for emphasizing this principle "is not to be sought in any requirement of a merely formal nature, but rather in an endeavour to invest the principle of causality with the authority of a law which holds good in the world of actual experience. One must above all avoid introducing into physical laws side by side with observable quantities, hypotheses which are purely fictitious in character, as *e.g.* the space of Newton's mechanics."

The bogey of causality need occasion no alarm. Had it been the guiding principle in physical theory little, if any, advance had ever been accomplished. The causal characters of scientific events are indeed generally expressed in terms of non-observable entities. The atom,

[2] Freundlich: *op. cit.* pp. 8, 25.

the molecule, the electron, the light-wave, the quantum, are co-ordinating and unifying influences in physical theory, but who has ever observed them? or who ever will? And if we are to be confined to purely observable quantities where will relativists find themselves, for to find a starting point at all we saw that we had to assume the existence of inertial systems far removed from all matter. We can never hope to observe the existence of such systems.

Although therefore it is true that when dealing with rotational phenomena we can observe only the difference between inertial and gravitational force, we are not guilty of any absurdity in keeping distinct the concepts of an axial and a central field; indeed, this conceptual distinction seems to me to afford the only possible justification for the assumption of the Minkowski invariant as a starting point for the general theory. The criterion which ought to guide us in this matter should be: "Which theory leads to the simplest description of the observed facts?" and in view of the general lack of mathematical knowledge there can be little doubt about the answer.

Einstein, however, will have none of this distinction. At this stage he brings in his principle of equivalence, according to which there is no difference between an inertial and a gravitational field. He supports this by the following consideration:—An inertial field originates, as we have seen, in the motional circumstances of the observer; it depends upon his perspective; it is therefore purely geometrical in its nature. Hence the path of a particle in such a field will be independent of the nature of the particle; it depends only on the geometry of the perspective. Now a purely gravitational field exhibits the same peculiarity, for we know that all bodies respond to gravitation in the same fashion, irrespective of their

nature.[3] Hence we are justified in ascribing exactly the same properties to the two kinds of field; we can draw no valid distinction between them; henceforward we must consider only one field in which inertia and gravitation play identical parts. The metrical properties of this generalized field will be euclidean in small domains because both inertial and gravitational fields can be transformed from such limited regions; but over large domains they will be non-euclidean because there will always be some residual field which cannot be transformed away.

I feel myself that the Einstein theory, regarded as a logical system, is at this point unsatisfactory. If inertial and gravitational fields were in every respect identical they should have identical properties; if they have not identical properties we can surely utilize that diversity in order to discriminate them. Einstein refuses to allow us to make the distinction within attainable regions of the manifold, yet he assumes that in the limit, *i.e.*, at an infinite distance from gravitating matter, the field is purely inertial and that the metric relations are there transformably euclidean. One may well ask, if the distinction is not valid within the solar system how can it become so at infinity? Moreover, the mathematical characteristic of a pure inertial field is the vanishing of the 20 Riemann symbols; if a pure inertial field and a pure gravitational field were identical in every way these symbols ought surely to vanish for the one as readily as for the other. But they do not vanish for a pure gravitational field. Einstein says that we cannot take the vanishing of those symbols as the general law of nature, as that would not allow of the existence of non-trans-

[3] I have seen a recent newspaper report of experiments described by Dr. Brush before the American Philosophical Society in April of this year. He claims to have established experimentally that gravitational and inertial mass are not identical. In the absence of details I can do no more than mention the report here.

formable fields which exist in the neighbourhood of gravitating matter; so that on his theory inertial and gravitational fields both are, and are not, the same. That position I find embarrassing logically. We can, indeed, make these fields approximately the same within limited regions, just as we can draw accurate maps on a plane of limited portions of the surface of the earth. As a useful approximation Einstein's theory is undoubtedly valid; but it is hardly a satisfactory method of meeting logical difficulties to keep pushing them out of sight.

Einstein seeks less stringent conditions by grouping the 20 symbols in a certain way so that the individual symbols may have values other than zero while the sum of each group vanishes. There is a very large number of ways of grouping these symbols together and one would like some logical guidance in making a selection from that plenitude of possibilities. But one looks in vain. The Einstein method of grouping seems to have been adopted because it looked the simplest way of doing it while satisfying the general principle of relativity. The result of that grouping he proposes as his fundamental law of gravitation. I cannot tell you what that law is, for it is not expressible in words. It is formulated in 10 differential equations; each equation involves 25 terms; these terms are built up out of 40 symbols, and each of these symbols is obtained by summing 12 terms. Even a professional mathematician may well stand aghast at this wealth of symbolism.

We may perhaps best grasp Einstein's law as a generalization of Newton's law of inertia, according to which a particle unacted upon by force moves with constant speed along a straight line path. Einstein's law is equivalent to a statement that "a material particle, uninfluenced by impacts with other particles moves along a geodesic within the space-time manifold." The

apparent shape of the geodesic depends upon the perspective in which it is viewed; and that perspective is determined by the motional circumstances of the observer, and by the presence of gravitating matter. But although we may interpret the law in such comparatively simple fashion, I do not know if many of us share the enthusiastic optimism of Schlick,[4] who says, that by taking this law, "we have considerably simplified the physical picture of the world, and consequently have now advanced another step in the theory of knowledge, by banishing gravitation, the last force acting at a distance, out of physics, and expressing all the laws underlying physical events solely by differential equations." Still, Einstein's law has been hailed with enthusiasm, not so much on account of its complexity, as by reason of its embracing within its scope phenomena which did not find a natural place within the Newtonian scheme.

[4] Schlick: *op. cit.* p. 64.

FOURTH LECTURE.

ON GENERAL RELATIVITY (*Continued*).

SECTION 7.—THE OLD AND THE NEW LAW OF GRAVITATION.

In regions where the field is weak Einstein's law degenerates into Newton's law, so that for most practical purposes the two lead to the same result. This circumstance seems to be misunderstood. Eddington,[1] in dealing with the Einstein law, says: "Whether it is (the general law of gravitation) or not can only be settled by experiment. In particular it must, in ordinary cases, reduce to something so near the Newtonian law that the remarkable confirmation of the latter by observation is accounted for." Schlick[2] says: "It is . . . . surprising that the new equations which have been obtained by such different means actually degenerate into the Newtonian formula for general mass-attraction." Einstein[3] himself says of it that "although fundamentally different from Newton's law (it) corresponded so exactly to the latter in the deductions derivable from it that only very few criteria were to be found in which the theory could be decisively tested by experiment." And again[4] "The fact that the equations deduced from the postulate of general relativity by purely mathematical processes give, to a first approximation, the Newtonian law and, to a second approximation, the motion of Mercury's perihelion, is a convincing proof that the theory is physically correct." One thus gets the idea that relativists look upon the two theories as being quite distinct in their physical bearing.

We must realise that any mathematical process no

---

[1] Eddington: S.T.G. p. 90.
[2] Schlick: *op. cit.* p. 63.
[3] Einstein: Nature: 106, p. 784, (1921).
[4] Quoted by Schlick: *op. cit.* p. 63.

matter how intricate it may be can lead only to new relations, and not to new physical concepts. We can liken it perhaps to the operation of power distribution. In a power station you have certain complicated mechanisms, and distributors issuing therefrom; by themselves and in themselves they are impotent. Put in energy, however, in the form of coal and air at one end, and you can withdraw energy in the form of light, or of heat, or of locomotion at the other. But the mechanism does not create the energy; its function is that of transformation and transmission.

If therefore we wish to compare the Einstein and the Newtonian theories, we must not lose ourselves in a maze of differential equations, but rather concentrate upon the bases upon which their theories are founded. The basis of the Newtonian view, as modified by Laplace and Poisson, is that some activity issues from gravitating matter, its intensity depending upon the mass of matter involved, and that this activity spreads throughout the neighbourhood according to the natural law of spread for space. On his view the mass of any portion of matter was constant, and the space was euclidean.

The Einstein theory is based upon :—

(i) The general condition termed the principle of relativity, that all magnitudes, mass, length, time, depend upon the perspective in which they are measured;

(ii) The condition that perspectives change with motional circumstances;

(iii) The principle of equivalence, namely, that gravitational fields produce a change of perspective just as motion does;

(iv) The special circumstance that such changes of perspective are appreciable only for large changes in relative velocity or for large differences of gravitational field.

But all these deal only with *changes* of perspective; they are all transformational. To get definite physical results we must assume some initial perspective to be known. For that starting point Einstein assumes the Minkowski perspective, and of that the Newtonian perspective is a particular case. What surprises me therefore is not that the Einstein law should give the Newtonian as a first approximation, but that it could ever have been expected to do otherwise.

SECTION 8.—EXPERIMENTAL TESTS OF EINSTEIN'S LAW.

Differences between the new law and the old can arise only from differences of perspective, whether due to relative motion or to gravitational fields; and, as these differences of perspective are seldom appreciable, we should not expect to find many opportunities of discriminating between them. Einstein, however, was able to suggest three different phenomena as likely to afford measurable discrepancies, and in two of these his predictions have been completely and triumphantly verified.

Of these phenomena, the first is the motion of the perihelion of Mercury. On the Newtonian theory a planet moving round a sun, and undisturbed by other forces would describe an elliptical path. In the solar system, however, such disturbing forces do arise owing to the influences of other planets; at any point of its orbit the net force of attraction upon the planet towards the sun is not exactly according to the inverse square law because bodies other than the sun also attract it. Such disturbing forces cause perturbations, that is to say, deviations from the elliptical path. One of these perturbations consists of a slow rotation of the axes of the orbit, so that the perihelion of the planet, or point of closest approach to the sun, moves along the orbit. In the

case of the planet Mercury—the innermost planet of the solar system—it is found from astronomical observations extending over long periods that its perihelion advances at the rate of 574″ of arc per century. But when the disturbing effects of all the other planets are computed according to the Newtonian theory, it is found that they account for an advance of only 532″ per century, leaving a balance of 42″ unaccounted for. For a long time it was thought that there must be another planet in the solar system whose presence would account for this additional term; but persistent astronomical investigation has refused to reveal it, and the discrepancy has remained unexplained.

On the Einstein theory the path of a planet will never be a perfect ellipse for two reasons :—(i) The mass of a planet is not constant, but depends upon its motion relative to the observer, and (ii) the metric relations in the manifold surrounding the sun are not euclidean, hence the law of spread in such region will not be the law of the inverse square. Einstein computes the equations to a geodesic in the gravitational field of the sun; he chooses the arbitrary constants involved so that the equations are in as close agreement as possible with the Newtonian orbit; when that is done it is found that there is a small residual term, which, on evaluation for Mercury, is found to give an advance of 43″ per century. The agreement is striking; the numerical concord issues from very general considerations, and is not based upon any *ad hoc* hypothesis. However much we may criticise Einstein's theory from the point of view of its logical difficulties, we must admit that he has accounted for a discrepancy that is otherwise hard to explain.

The second crucial phenomenon involves the behaviour of a ray of light in passing through a gravitational field. Suppose a ray of light is coming towards us from a

distant star. As long as it is passing through the void we see it in a constant perspective; according to the special theory of relativity it travels towards us with constant speed in what we term a straight line. Suppose now that a focus of gravitational attraction intervenes between us and the star. According to the principle of equivalence we may replace the field due to that focus by the corresponding field of acceleration. As the ray passes through the field we must give it at every point the acceleration characteristic of that point. You can see that the effect of that addition will be a bending of the path of the ray towards the centre of attraction. Nor is that the only effect. As the ray deviates, our perspective changes owing to the change in its motion relative to us; and the effect of this change of perspective is to occasion a further deviation in the same direction. The deflection therefore consists of two parts, the one accounted for in terms of the principle of equivalence, the other due to the change of perspective occasioned by the first constituent. As the velocity of light is extremely large there would be no hope of observing the predicted effect except in intense gravitational fields, because otherwise the light ray would shoot through the field before it had any chance of effecting a deviation. The largest gravitational field that we can experiment with is that of the sun, and even then a light ray must pass extremely close to it in order to give a deflection at all measurable. Since the glare of the sun renders completely invisible any stars behind it, the only chance of observing the predicted effect would be during a total eclipse. On the Einstein theory the deflection for grazing incidence should be $1''{\cdot}75$; if light were like a material particle and subject to gravitational acceleration, then on the Newtonian theory the deflection at the sun's limb should be only the first constituent of the Einstein effect, about $0''{\cdot}87$; and if, as had generally been

believed, light were not subject to gravitation at all the deflection should be zero.

Two expeditions were sent from England to view the eclipse of May 29th, 1919—one to Brazil, and the other to West Africa. The results obtained by these expeditions were once more in striking confirmation of Einstein's theoretical predictions. The average value for the shift at the sun's periphery was found to be $1''\cdot98$ by the Brazil party, and $1''\cdot61$ by the West African party. These figures show pretty conclusively that the deflection does exist, and that its value is larger than that given by the Newtonian law of attraction.

Attempts have been made to explain this bending of the light beam, without having recourse to the Einstein theory. It has been suggested that the fall of temperature in the Earth's atmosphere at the place of observation as the moon cuts off the sun's rays might occasion abnormal atmospheric refraction, and thus give rise to the observed deviation; but calculation shows that the observed shift could not be accounted for in that way; and, in any case, it so happened that owing to unusual early morning cloud the fall of temperature at one of the places of observation was only a few degrees Fahrenheit. The question has also been raised whether the observed bending might not be due to refraction by the sun's atmosphere; but that would require an impossibly high density of the sun's atmosphere.

An aether, suitably modified in the neighbourhood of the sun, could, of course, account for the observed displacements, but one shrinks from such *ad hoc* hypotheses. One naturally prefers the theory from which the phenomenon was predicted.

The third test is rather more technical than the other two. I presume you are all familiar with the Doppler effect, although you may not know it by that name. If

you are in a railway train and a whistling engine passes you, you will notice a sudden drop in the pitch of the note as the engine passes; the faster the relative motion between the trains the larger is the drop. With light a similar effect is experienced, the colour of the light playing the same part as the pitch of the note; but in the optical case the effect is small for attainable astronomical velocities and can be detected only by refined spectroscopical measurements. If the observer and the source of light are approaching, the colour moves a little towards the violet end of the spectrum; if they are separating the colour moves a little towards the red end.

Now consider light originating in the vibrations within an atom in a lamp in our laboratory and light arising from the vibrations within an identical atom in the sun. If there were no difference between the circumstances of vibration in the two cases we should find the lights to be of exactly the same wave-length, that is to say of the same colour. But the atom in the sun differs from our terrestrial atom inasmuch as it is in a much stronger gravitational field. According to Einstein's principle of equivalence a gravitational field is identical with a field of acceleration, so that the attractive force towards the centre of the sun can be replaced for the sun's atom by the equivalent accelerated motion. That would mean a recession from the earth and the effect of such recession should be evidenced by a shift of colour towards the red end of the spectrum compared with the light from an identical atom on the earth. According to the theory such shift should always be towards the red end; it should be proportional to the wave-length of the colour under investigation, and should be independent of the intensity of the light. Anxious search has been made for experimental verification of this prediction. Most of the spectrum lines examined do show differences of position

according as the source is on the earth or in the sun, but the shifts are sometimes towards the red, sometimes towards the violet; they do depend upon the intensity, and they are not proportional to the wave-length.

Although this third prediction of Einstein has not yet found experimental verification, it must not be thought that his theory must therefore be scrapped. It may well be that the effect does exist, but that it is masked by other effects. Surface currents and pressure gradients in the solar atmosphere would also cause shifts of the spectral lines, and astronomers' knowledge of solar meteorology is not yet sufficiently accurate or sufficiently extensive to enable us to estimate separately these masking effects.

The most recent work upon the question has been carried out by Perot,[5] working with lines in the magnesium spectrum, and by Buisson and Fabry,[6] who used lines in the spectrum of iron. These scientists have come to the conclusion that the pressure in the reversing layer of the sun's atmosphere is negligible in so far as its influence upon these two spectra is concerned; and they state that if the solar spectra are compared with terrestrial spectra obtained under the same pressure conditions—that is, in a vacuum, instead of under atmospheric pressure— then the shift towards the red is of the magnitude predicted by Einstein.

In the meantime a lot more work must be done upon solar spectroscopy before it can be asserted that the third Einstein prediction is verified or not. It is somewhat unfortunate that this test should fail to carry conviction, because the other two, which have been verified, could be accounted for on a suitably modified aether theory, but the third could not. In the meantime, therefore, the only thing to do is to await further experimental evidence on the point.

---

[5] C.R. 172. No. 10. p. 578, (1921).
[6] C.R. 172. No. 17. p. 1020, (1921).

SECTION 9.—REVIEW OF THE MAIN ASPECTS OF THE THEORY OF RELATIVITY.

Before proceeding to consider Einstein's cosmology, we might usefully review the main features of the theories that we have been examining.

In formulating any theory of physical knowledge, the accounts of all possible observers must be synthesized, otherwise the theory will be but partial. We assume that other observers are human beings, like ourselves, endowed with the same fundamental mental structure, and with the same basic order-intuitions. We find from experience that their descriptions of observed events do not always agree with ours, and the correlation of such divergent descriptions we termed a *transformation* from one point of view to the other. Permanences unaffected by changes of point of view we called *invariants*.

As long as we are dealing only with observers who are at rest relatively to one another we assume that time is invariant, as is also the distance between fixed points upon a rigid body; all such observers ascribe the same value to a given time interval or to a given distance interval. These invariants yield the ordinary heterogeneous manifold of local space and of separate time which forms the customary system of reference into which we fit our experiences.

When we try next to take into account the descriptions of observers who are in motion relatively to one another we first simplify matters by stipulating that their relative velocities shall be uniform, that is to say, that they shall take place at constant speed in straight lines. It would be natural to endeavour to make the same invariants still serve our purpose of describing objective reality as in the case of stationary observers. We find that the principles of motion and of congruence still hold good, but velocities vary according to the observer's point of

view. Now as long as we are dealing with bodies moving on the surface of the earth, we can express their velocities relatively to one another in terms of their motion with respect to the earth; but when we deal with other motions, we are faced with the difficulty of interpreting motion through space. It had hitherto been thought that the aether gave a system of reference in space to which such motions could be referred; but the Michelson-Morley experiment showed that it is impossible to detect such motion. There is nothing in space to reveal motion through it, and, on the Einstein view, a consequence of that ignorance is the postulate of the invariance of the velocity of light for all observers. Obviously, then, the invariants of the Newtonian transformations, according to which velocities do vary from one observer to another will no longer suffice, and consequently time-intervals and distance-intervals cannot be constant for all observers.

The transformations of the special theory of relativity which take the place of the Newtonian transformations, are the Lorentz group. They render the manifold homogeneous; objective reality is retained within the manifold by the discovery of the Minkowski invariant for the interval between two neighbouring events. Changes of perspective now involve changes not only of spatial but also of temporal points of view. The Minkowski invariant gives the interval between two events in a region of the manifold far removed from the influence of gravitating matter as measured by any uniformly moving observer; that interval is an element of a world-line, or geodetic line within such region of the manifold, and for such observers these geodesics are straight lines. The Minkowski invariant is concerned only with objective reality underlying phenomena which do not involve accelerations or gravitational fields; it can therefore cover only rectilinear motions and optical phenomena—a com-

paratively restricted portion of the physical phenomena with which we have to deal.

Hitherto we have been synthesizing the descriptions of a special class of observers only; the next step towards generalization is to obtain the transformations correlating the points of view of all observers no matter how they may be moving relatively to one another and no matter how they may be situated with respect to gravitating matter. This generalization is based upon the relativity of all motions and upon the principle of equivalence. The Minkowski invariant no longer suffices; there is no longer a fixed and definite interval between two events, it depends upon the perspective from which the events are viewed; the only objective knowledge left is that events happen, such happenings being conditioned by the intersection of geodesics within the manifold. If that were all we knew we could effect no synthesis; we should just have to take events as we found them. But it is assumed that in a region far removed from matter a fundamental observer can be found for whom all geodesics are straight lines. Einstein interprets this by saying that in such region the space-time manifold is "flat." In the neighbourhood of gravitating matter no such fundamental observer can be found. Einstein says that since in such regions all geodesics cannot be made straight simultaneously the manifold is there "warped." According to the principle of equivalence, within a limited region which does not actually contain matter the geodesics can be made straight. But singularities are encountered—points at which the geodesics are, as it were, blotted out; such a singularity we call a particle of matter.

The special theory of relativity welded space and time into a homogeneous manifold; now matter as an independent entity must also go, for a particle of matter

is but a singularity in the space-time manifold. The geodesics within the manifold form natural tracks along which events must take place; they move along these tracks whether they be slow moving events like Mercury in its orbit, or fast moving events like the passage of a light beam from a distant star. The idea of force is also banished from physics; it is no longer wanted. Planets revolve about the sun not because they are compelled by the force of gravitation to deviate from straight line paths, but because they simply follow the natural tracks in the gravitational field of the sun. The field of reality is reduced to the space-time manifold, and everything else is geometry. Einstein might well adopt as his text the aphorism of Galileo: "The Book of Nature is written in characters of Geometry."

SECTION 10.—EINSTEIN'S COSMOLOGY.

It is at this stage that cosmological considerations intrude.

You have seen the Foucault pendulum, and have considered the usual interpretation of the phenomenon as revealing the rotation of the earth. According to the general principle of relativity we may just as well say that the earth is not rotating, but that it is surrounded by an inertial field, and that this field causes the plane of swing to rotate. This inertial field increases in intensity as we recede from the earth's surface, but how are we to account for it? It cannot be due to terrestrial circumstances, for the earth is no longer rotating; it must originate in relative motion; it must therefore be caused by bodies outside the earth. If our earth were the only world in the universe, there could, according to this view, be no equatorial bulge, no rotation of a Foucault pendulum. Can we really believe that the stars in their

courses conspire to twist the plane of swing of this little instrument?

Whether we are willing to accept this explanation or not, calculation shows that the known motions of the stellar universe could not produce the centrifugal field of the earth. We find ourselves, therefore, in this dilemma: Either we must ascribe absolute significance to rotations, obtaining for that purpose some absolute frame of reference, or we must explore the outermost recesses of the universe, and there find the explanation either in the presence of hitherto unsuspected vast quantities of world-matter, or in the geometrical structure of the cosmos itself.

Eddington holds that the geodesics or natural tracks form an absolute feature of the world with reference to which rotation becomes absolute. For myself, I cannot see how this is possible. It is true that for a super-observer existing in five dimensions the geodesics would be absolute, just as we have drawn in two dimensions the absolute world-lines of events in a one-dimensional world. But to assert that for us, within the manifold, absolute geodesics exist, seems to run counter to the whole doctrine of relativity; we would then know what world-lines were straight, and what lines were not straight, and we could therefore differentiate absolutely between uniform and non-uniform motions. That is not possible, for as we have seen, geodetic structure within any region depends upon two things: (i) the intrinsic nature of that region, and (ii) the perspective from which the events are viewed. If observed geodesics do not seem straight to us, we may attribute this to a rotating system of reference, or we may just as readily attribute it to the intrinsic nature of the manifold itself, that is to say to the presence of gravitating matter. Rotation relative to geodetic structure seems to be meaningless.

Einstein, himself, differs on this point from Eddington, and he prefers to explore the universe in search of the cause of the centrifugal field.

He first mentions certain difficulties which he finds in Newton's cosmology. On the Newtonian view space is infinite, and time eternal. But although space is infinite, it cannot be occupied by matter everywhere. The Newtonian theory of gravitation necessitates a cosmos with a more or less definite centre. The average density of matter in the universe is greatest in the neighbourhood of that centre; and as we recede from it, the average density diminishes, the worlds we encounter get fewer, until ultimately we reach the void, unoccupied and desolate. The average density of matter throughout all space is vanishingly small, for the mass of matter in the universe is finite while the space in which it is contained is infinite. Heat, light, and other radiant energies issuing from the suns of the universe travel onwards and outwards, and are lost for all time in the immensities of space; return is impossible, and thus a perpetual waste of energy is proceeding. In the course of time our sun will have radiated away all his light and heat; other suns will do the same, so that the universe must necessarily die.

Einstein calls it a distasteful conception that the material universe ought to possess a centre, and considers equally unsatisfactory the inevitable impoverishment of the universe by radiation into space. The idea of a finite universe with a definite centre adventuring through space and time seems to me easier of comprehension, and therefore less distasteful, than an unending realm of worlds distributed throughout all space, so that no matter how far you explore in any direction, new suns and new planets will be encountered. Not only is the concept less distasteful, but it seems to be substantiated by our present knowledge of stellar cosmology. Astronomers tell

us that the known stars form an approximately lenticular system distributed more or less symmetrically about the galactic plane, the density of the distribution diminishing as we recede from that plane.

And as for the de-energizing of the universe, what more natural concept could we form? The death of worlds should be no more unexpected than the death of peoples; change, decay and death are ultimate unavoidable occurrences.

Such considerations, therefore, hardly seem to afford adequate reasons for interfering with the Newtonian cosmology. But the discovery of non-euclidean relations within the space-time manifold of the general theory of relativity raises the question of the possibility of the manifold being unlimited in every direction and yet not infinite.

Einstein's first idea was that the manifold is flat everywhere except in the neighbourhood of matter, where warps or hummocks occur; just as if we were to place a number of marbles upon the table and cover them with a sheet, the sheet would be flat everywhere except for the isolated hummocks caused by the marbles. Such structure he calls quasi-euclidean, and it would be rendered more or less probable by the consideration that the local variations in the intrinsic structure of the manifold are hardly appreciable even in the immediate neighbourhood of the large mass of the sun. But this supposed structure leaves the cosmological problem in the same position as Newton's theory. The average density of matter throughout space would still be zero, so that the universe would still be a finite aggregate of matter in an infinite space.

On the relativity theory this conception is less acceptable than on the Newtonian view. We have seen that if the earth is to be regarded as being at rest we must probe the outermost recesses of space to find a cause for

its inertial field. The further off we go the larger becomes the centrifugal force to be accounted for; it increases without limit. But on the relativity theory inertia is due to the presence of other bodies, and these cannot be found in emptiness; so that no explanation can issue from the void.

We ask ourselves can anything exist in empty space infinitely removed from all matter. No physical events can happen there, for a physical event is revealed to us through the agency of a material object, and there is no matter there. Since there are no events, there can be no intervals. Since there are no intervals, there can be no geodesics, which are built up out of intervals. Were there no intervals anywhere our manifold would collapse, shrink away to nothing; whatever causes intervals must therefore provide the extension of the manifold. If the manifold were infinite in every direction, as supposed on the quasi-euclidean structure, it is hard to conceive how such extension could be caused by the finite field of events within it, just as when we wish to stretch a sheet of rubber we must apply forces outside it. For such reasons Einstein rejects the quasi-euclidean structure.

You may remember that when we were discussing the geometry of a spherical surface, we saw that what two-dimensional beings upon such a surface would call a straight line, would be for us a great circle on the surface. Their lines would be unbounded, inasmuch as they would have neither beginning nor end, nevertheless they would be of finite length. As they drew circles of larger and larger radii on their surface, the circumferences would be initially π times the diameter but the circumferences would increase less rapidly than the diameter until a maximum circumference would be encountered, and then the circles would begin closing up again as the diameter still increased, and would ultimately shrink up to a

point—the antipodal point. Their surface would be unbounded in any and every direction, and yet it would be of finite area, or of limited extent. If only the radius of their sphere were large enough compared with the scale of their measure operations, they could never suspect that their surface were anything other than a flat surface, or their straight lines other than infinite as well as unbounded.

Einstein suggests that it may well be that our space is also unbounded in every direction and yet limited in total content. By analogy, we might call such space spherical space, but that expression must not be interpreted as meaning shaped liked a sphere. If from any point in such a space we start describing spheres, the surfaces of these spheres would be initially 4 π times the square of the radius. But as they became larger, the ratio of the surface to the square of the radius would become smaller; we should reach a sphere of maximum surface; on taking still larger radii, the surface would begin to diminish, until for a certain radius the sphere would shrink to a point. We could describe no more spheres then; our space would be full. Although its total content is limited, nevertheless it is not bounded in any direction. Attempts have been made to measure the "space constant," but all we can say about it is that it is extremely large compared with the radius of the earth's orbit; and this accounts for the fact that even our largest scale astronomical measurements are sensibly euclidean in their relations. Macroscopically the Einstein space is spherical; but when small portions of it are examined, local warps and strains are encountered wherever matter is present; just as we may say that a well-shaped orange is spherical, although closer examination reveals irregularities and rugosities of surface. This limitation of space gets us out of the difficulty regarding the centri-

fugal field. We follow the inertial force further and further afield, hoping to pin it against the boundaries of space, and make it explain itself; but we are chasing a will-o'-the-wisp. As we follow it into space it at first increases; then it reaches a maximum; as we proceed further, the force diminishes, and at the antipodal point it vanishes altogether. There is then no need for any explanation because there is nothing left to explain. The centrifugal field, like everything else, originates in the geometrical structure of the world.

Since space is now bounded the cause of its extension may well be contained within it, just as we can inflate a balloon by internal pressure.

Up to the present we have been considering only the boundedness of space, that is to say, the finite content of an instantaneous section of the manifold. But we are also interested in time. In Einstein's cosmic manifold the time axis is straight, so that time is for him unending and infinite. His manifold is, as it were, a four-dimensional cylinder, bounded in its three space dimensions, but extending to infinity in time. It is not an easy thing to imagine. Moreover, it seems to be inconsistent with his own relativity theory inasmuch as it brings back once more an absolute distinction between space and time—space is unbounded but finite, time is unbounded and infinite.

The theory leads to some interesting speculations. Since space is bounded no radiant energy can escape from the universe. The light and heat issuing from a sun or star travel on, and if there is no absorption in space, are brought to a focus once more at the antipodal point many years later. It may well be, then, that a number of the stars we ponder over nightly are only ghosts, condemned to eternal re-formations of radiant foci after their voyages round the universe.

The Dutch astronomer, de Sitter, has suggested a different cosmology. He curves the time axis as well as the space axes, so that his cosmos is completely bounded both in space and time. The idea that time is closed and not infinite is a favourite conception of the mystical poets, as, for instance, in these opening lines of Henry Vaughan's beautiful poem, "The World":—

"I saw Eternity the other night
Like a great Ring of pure and endless light,
All calm as it was bright."

At first sight it looks as if bounded time would put us in the difficult position that events would have to repeat themselves; the history of the universe would then form a closed series, and history would repeat itself exactly and inevitably. In de Sitter's cosmology, however, we are saved this embarrassment by the circumstance that as we look further and further afield in the manifold, time seems to run slower and slower, until, when we view a point half-way to the antipodes, time comes to a standstill. At that point there is an impenetrable barrier; light, heat and radiant energy are there brought to rest; they never pass right round the world, so that there can be no repetition.

These imaginative cosmologies are interesting and diverting, but they are very vague and elusive. I look upon them as somewhat nebulous excrescences upon the scientific theory of relativity; there is, in the meantime, little hope of being able to put them to any experimental test.

Section 11.—"Cui Bono?"

In conclusion, I may, perhaps, attempt to answer the question that some of you have been very anxious to ask: "What is the use of Relativity to anyone?" For myself, I never worry whether a thing is useful or not; if only it

is interesting it justifies itself; but I quite recognize the existence of other standards.

Relativity might prove of use in industry. A point in the special theory which I did not mention is the identity of inertia and energy. If this is the case then every material body contains a hitherto undreamt of store of energy, that might possibly prove available in the distant future.

Relativity may prove of use to astronomers. The computation of astronomical orbits is at present a laborious process; it may be that mathematical methods will be worked out for the computation of a geodesic in the manifold that will lighten appreciably the present labour of astronomers.

Relativity has already proved of utility in philosophy, for it has necessitated a re-examination of the fundamental concepts of space and time and a more rigorous statement of the bases upon which the ordinary views are founded.

Relativity has proved of utility in Science generally, for it has made us think; it has brought home to us the limitations of the human understanding, and has made us more critical of any scientific dogmatism.

Those of us who endeavour to explore the realms of abstract thought find there a veritable fourth dimension; it gives us a "refuge from immediate fact," an "escape from the tyranny of time"; if we follow a track within that region we find it indeed

> "A road where restful Time forgets
> His weary thoughts, and wild regrets."

Of such roads one of the most fascinating is that discovered by Einstein, and his pioneering labours will shed undying lustre upon himself and on the race from which he springs.

INDEX.

HORTORS LIMITED,
LITHOGRAPHERS & PRINTERS,
JOHANNESBURG.